Entropies et Radiotechnique

Jean-François Bercher

Université Paris-Est
Marne-la-Vallée

ESIEE-Paris

Mémoire d'habilitation à diriger des recherches

Entropies et radiotechnique

Jean-François Bercher

Équipe Signal et Communications
Laboratoire d'Informatique Gaspard Monge
Université Paris-Est
et
Département Signaux et Télécoms
ESIEE, Université Paris-Est.

Présenté le 1er décembre 2009 devant le jury constitué de :

M	Jean-Marc	BROSSIER	Professeur INPG	Rapporteur
M.	François	CHAPEAU-BLONDEAU	Professeur Université d'Angers	Rapporteur
M.	Ali	MOHAMMAD-DJAFARI	Directeur de Recherches CNRS	Rapporteur
Mme	Geneviève	BAUDOIN	Professeur ESIEE	Examinatrice
M.	Patrice	GAMAND	Directeur centre R&D NXP Caen	Examinateur
M.	Jean-Christophe	PESQUET	Professeur Université Paris-Est	Examinateur
M.	Christophe	VIGNAT	Professeur Université Paris-Sud	Examinateur

Table des matières

À Rosalind Franklin

Chapitre 0

Préambule

Ce document présente une synthèse de mes contributions dans le domaine de la recherche depuis la soutenance de ma thèse de doctorat en 1995. L'activité rapportée ici est celle d'un enseignant-chercheur, et s'est donc déroulée en parallèle d'une activité d'enseignement comprenant également certaines charges pédagogiques, d'organisation ou d'ordre administratif. Je suis enseignant-chercheur à l'ESIEE depuis 1996 (professeur assistant, puis professeur associé (2002) au sens anglo-saxon). L'ESIEE est la grande école technique de la Chambre de Commerce et d'Industrie de Paris (la CCIP est un établissement public, l'école étant sous tutelle du ministère de l'industrie).

Lors de la préparation de ma thèse de doctorat, je m'étais intéressé à une méthode de résolution problèmes inverses reposant sur une approche à maximum d'entropie. C'est assez naturellement que j'ai continué à investir sur le thème du traitement et la théorie de l'information, et notamment sur l'étude de mesures d'information ou de complexité pour l'analyse, le filtrage et la restauration de signaux et images, avec des applications aux télécommunications et à la physique statistique. D'un autre côté, je n'ai pas complètement oublié que j'ai débuté mes études supérieures dans le domaine de l'électronique (bien qu'assez fortuitement). C'est par l'interaction avec les collègues de l'ESIEE, et en coordination avec des activités d'enseignement, que j'ai été amené à m'intéresser à certains problèmes d'architecture des émetteurs-récepteurs radio. Cette activité, que j'appelle "Radiotechnique" (en écho au titre du livre fascinant [1]), s'est beaucoup développée ces dernières années.

Ces deux volets de mes recherches sont menés tantôt alternativement, parfois conjointement. Les aspects concernant le traitement de l'information et ses applications sont présentés dans les chapitres 2 à 4, quand les aspects sur la radiotechnique sont décrits dans les chapitres 5 et 6. Il aurait été possible, à l'occasion de ce mémoire, de relier ces deux thématiques ; par exemple par le biais du codage, ou de l'égalisation. En réalité, ces deux axes sont indépendants et je ne souhaite pas afficher ici de montage artificiel. Il est bien possible que ces deux axes puissent se rejoindre ou échanger dans le futur. Le présent mémoire n'est ni une finalité, ni un achèvement, bien heureusement, et en recherche comme ailleurs, la cohérence s'imposera par ses propres qualités.

L'exercice de rédaction et soumission d'un manuscrit d'HdR permet de faire le point, et de recueillir l'appréciation d'autrui (mais on sait bien, sauf peut-être dans les sommets de l'organisation de l'État, que l'évaluation de la recherche est permanente). L'exercice constitue un jalon dans le devenir d'un enseignant chercheur. Incidemment, le grade d'HdR est important pour les autres : les dossiers de reconnaissance, de demande de crédit, les réponses aux appels d'offre, les évaluations d'équipe ou d'établissement comptabilisent l'HdR ou le nombre d'HdR comme un indicateur de qualité. C'est également un jalon en termes personnels, parce qu'il permet sans doute d'aller plus loin, en gagnant un peu en stature (encore le regard des autres) et le droit de développer des recherches autonomes. En réalité, cela fait bien longtemps que j'assume des recherches personnelles. Celles-ci ne sont toutefois pas sans interaction avec autrui. En particulier C. Vignat et C. Berland auront contribué largement aux thématiques et résultats présentés ici. La possibilité de diriger en direct des travaux sur le thème de la Radiotechnique permettra de développer plus avant ces aspects. La possibilité de pouvoir rechercher des étudiants pour des travaux sur les thèmes

plus amont, liés aux aspects entropiques est également fort intéressante.

Ma démarche de recherche est de type spagyriste – l'expression est de [2], et consiste à extraire, assembler, en vue de construire autre chose. Je conçois et profite de l'activité de recherche pour continuer à apprendre, et pas simplement à produire.

En rapport avec la dédicace en début de ce mémoire, je rappellerai les grands principes, qui selon J. Watson (co-découvreur de la structure de l'ADN), devraient conduire une carrière de chercheur. Si ceux-ci peuvent être un peu critiquables (notamment en raison de la personnalité de leur auteur), ils me semblent un bon guide (et une bonne motivation au présent manuscrit) pour aller vers une recherche originale et de qualité. Les voici : (i) tuer le père (au sens figuré, s'entend), le quitter et mener des recherches hors de son haut patronage (ii) ne fréquenter que des gens d'un niveau scientifique supérieur, ou au pire égal, de manière à profiter d'une dynamique et progresser (iii) refuser de faire ce que l'on n'a pas envie de faire (iv) trouver une position (au sens anglo-saxon du terme), ou se mettre dans une situation telle qu'on puisse mettre en application les points précédents.

La suite du document est organisée comme suit. Les pages suivantes présentent un bref curriculum-vitæ, quelques mots sur mon activité d'enseignement, sur la formation doctorale, l'encadrement de stages et thèses. Dans le premier chapitre, je présente rapidement un bilan de mes différentes contributions, le chapitre se terminant en page 13 par la liste des publications. Les cinq chapitres suivants reprennent une grande partie des résultats que j'ai pu obtenir, en s'appuyant sur les articles placés en annexe. J'essaie de donner l'essentiel du contexte, de la démarche et des résultats concernant chacun des travaux présentés. Lorsque les articles ne sont pas annexés, je donne parfois un peu plus de détails et d'illustration. Bien entendu et comme il se doit, le mémoire se termine avec quelques perspectives.

Montry, le 16 avril 2009.

[0] Le mémoire a été rédigé début avril 2009. Seules quelques mises à jour ponctuelles et la prise en compte de suggestions des rapporteurs ont été ajouées depuis lors.

BREF CURRICULUM VITÆ

Nom patronymique : *BERCHER*

Prénom : *Jean-François*

Date et lieu de naissance : *12 août 1965 à Marseille*

Nationalité : *Française*

Situation de famille : *Marié, trois enfants.*

Fonctions actuelles : *Enseignant-chercheur à l'École Supérieure d'Ingénieurs en Électrotechnique et Électronique (ESIEE).*
Cité Descartes, BP 99, 93 162 NOISY LE GRAND *Cedex*
Tél. 01-45-92-65-15
email : jf.bercher@esiee.fr
et Laboratoire d'Informatique Gaspard Monge, LIGM, UMR 8049, Équipe Signal et Communications

Titres universitaires

Doctorat de l'Université Paris-Sud (UPS),

Spécialité : « Automatique et traitement du signal »
Intitulé : « Développement de critères de nature entropique pour la résolution de problèmes inverses linéaires »
Directeur de thèse : M. Guy DEMOMENT, professeur à l'Université de Paris-Sud,
Date et lieu de soutenance : le 02 février 1995 à l'École Supérieure d'Électricité,
Mention : très honorable, avec les félicitations du jury.

Diplôme d'Études Approfondies (DEA) en « Signal Image Parole » obtenu mention Très Bien en septembre 1990 à l'Institut National Polytechnique de Grenoble (INPG).

Diplôme d'Ingénieur Électricien spécialité Traitement du Signal obtenu mention Bien en juin 1990 à l'École Nationale Supérieure d'Ingénieurs Électriciens de Grenoble (ENSIEG).

Maîtrise des Sciences et Techniques en Génie médical et Instrumentation obtenue en septembre 1988 mention Bien à l'Université Paris XII.

Diplôme Universitaire en Technologie (DUT) option Électronique et informatique industrielle obtenu en Juillet 1986 à l'IUT de Cachan.

Activités professionnelles

– septembre 2007-septembre 2008 - congé de perfectionnement au laboratoire des signaux et systèmes (Orsay) et dans l'équipe Signal et Communication du LabInfo (IGM-Marne la Vallée)
– depuis septembre 1996 : enseignant-chercheur à l'ESIEE (départements "signaux et télécoms"et "modélisation et simulation"). Ancien responsable de l'option Traitement du signal et Télécommunications, membre élu de la commission d'évaluation (commission de spécialistes)
– Attaché temporaire d'enseignement et de recherche - institut Galilée, Paris 13 (1995-96), université Paris-Sud (1994-95).

Activité de recherche

Depuis 2008, je suis membre de l'équipe "Signal et Communications" de LabInfo-IGM, UMR 8049. Auparavant, j'étais rattaché au thème "radiocommunications numériques" de l'équipe Esycom, EA 2552.
Les recherches que je développe comprennent deux axes :

1. le traitement et la théorie de l'information, et notamment l'étude de mesures d'information ou de complexité, appliqué à l'analyse, le filtrage et la restauration de signaux et images, avec des applications aux télécommunications et à la physique statistique.
2. La radiotechnique (statistique) : analyse statistique et calcul de performances pour les émetteurs-récepteurs numériques, développement et caractérisation statistique d'algorithmes de correction, calibration ou contrôle.

Les principaux résultats obtenus sur ces deux thèmes sont décrits dans le chapitre 1 et dans la suite du mémoire.

Formation doctorale

– Intervention durant plusieurs années dans le DEA IFA (Informatique Fondamentale et Applications) pour un cours de traitement et compression d'images.
– J'ai participé récemment à l'aménagement et la proposition d'une nouvelle spécialité, "Image Signal et Son (SIS)" dans le cadre de la filière Image du Master d'informatique.

Encadrement

Master et Post-Doc
– PostDoc 6 mois en 2003 - Mlle Pavla Holejšovskà - méthodes de reconstruction entropiques ; 3 communications
– 2007 : co-encadrement, avec A. Chevreuil, du stage de dernière année de l'ESIEE de Leire Azpilicueta et Pascal Vallet (ce dernier poursuit actuellement en thèse dans l'équipe sous la direction de P. Loubaton). Sujet : estimation d'un signal d'otoémission pour un procédé de diagnostic de la maladie de Menière. Financement par le contrat DREAMM.
– 2008 : co-encadrement, avec A. Djafari, du stage de Master de M. Boris Pougaza (au laboratoire des Signaux et Systèmes) - *Using the notion of Copula in tomography*.

Thèses
– Thèse Cifre de Cyril Joubert coencadrée avec G. Baudoin (directrice de thèse), 2004-200 ?, "contribution à l'analyse d'une boucle à verrouillage de phase (PLL) à traitement de signal numérique dans les émetteurs radio". 3 communications [C22, C26, C27] ont été présentées. Embauché par ST Microélectronics, Cyril se fait attendre pour soutenir effectivement sa thèse. Les résultats sont présentés en section 5.2.
– Thèse Cifre de Pierre Guillot, coencadrée avec C. Berland (directrice de thèse), 2006-200x, "Architecture des systèmes de radiocommunication à références d'horloges hautes fréquences : application des résonateurs BAW à la génération de fréquence de référence dans les systèmes de communication mobile", 2 communications [C31, C34] ont été présentées et un article accepté récemment [R14]. Pierre a obtenu récemment un « Best Student Paper » pour la communication [C34]. Cette thématique et les résultats sont présentés en 5.3.
– Thèse Cifre de Mlle Amandine Lesellier, débutée en octobre 2008, sur le sujet "contribution à l'étude des architectures de récepteurs large bande multicanaux". Pour la direction de cette thèse (coencadrée avec O. Venard), j'ai obtenu l'accord du conseil scientifique de l'université Paris-Est en septembre 2008.

Participation à d'autres encadrements

Je citerai également ici deux autres thèses pour lesquelles j'ai eu des contributions certaines à l'encadrement et aux résultats obtenus :

– Thèse d'Isabelle Jamin-Hibon "Analyse et conception d'une architecture numérique innovante d'émetteur universel pour radiocommunication numérique" (direction C. Berland et M. Villegas), soutenue en novembre 2006. Un article de revue [R6] a été publié et un brevet [V2] déposé. Je partage avec Corinne Berland l'idée initiale qui sous-tend le travail de thèse, a été déposé comme brevet et a fait l'objet de l'article de revue discuté au paragraphe 5.1 et annexé en page 166.
– Thèse d'Antoine Diet, "Etude des architectures EER pour les émetteurs de radiocommunications numériques 3G et au-delà" (direction C. Berland et M. Villegas), soutenue en mars 2005. J'ai ici contribué aux recherches sur l'évaluation, par méthode de Monte-Carlo, des performances obtenues en présence de défauts d'alignement. Ceci est présenté dans le paragraphe 6.1, a fait l'objet de la communication [C20] et du chapitre 5 de la thèse d'Antoine Diet.

Relecture

J'ai effectué ou effectue des relectures d'articles pour IEEE transactions on Signal Processing, Signal Processing (version Elsevier), Traitement du signal, Physics Letters A, Information Sciences, Electronics Letters

Activité d'enseignement

À l'ESIEE, j'interviens pour des cours, travaux dirigés et pratiques dans les 3 années du cycle ingénieur, dans le domaine du traitement du signal et des télécommunications. Je suis également intervenu dans l'école de techniciens supérieurs (à Bac +3), ESTE, appartenant au groupe ESIEE. ainsi que dans le programme N+I. J'ai fait parti du comité de pilotage de l'ISBS, Institut Supérieur des Biosciences de Paris, formation créée par la CCIP et en cohabilitation entre l'université Paris 12 et l'ESIEE.
Durant trois années, j'ai été responsable de la majeure télécommunications et traitement du signal, spécialisation couvrant les deux dernières années du cycle ingénieur. J'ai également été co-responsable du bloc électronique de tronc commun.

J'assure actuellement et ai assuré la responsabilité de quelques unités d'enseignement (programme, coordination, vacataires, etc).
– unité EL201 bases d'électronique et de traitement du signal (2^{e} année — depuis septembre 2000),
– unité EL301 filtrage numérique, signaux aléatoires (3^{e} année — co-responsable, responsabilité Gsignal 1997-2001),
– bloc traitement du signal et dsp, devenu Signaux et Systèmes 1, SIS1, (4^{e} année, majeure « systèmes embarqués »), depuis 2000 (au moins),
– bloc théorie du signal SIG1 (4^{e} année, majeure « télécommunications et traitement du signal »), depuis 2004,
– unité techniques avancées de communications numériques (5^{e} année) jusqu'à 2002,
– unité traitement satellite et radar (5^{e} année), entre 2002 et 2003,
– unité égalisation et synchronisation (5^{e} année), depuis 2005
– unité traitement du signal (1re année ISBS) depuis 2004
– unité théorie de l'information et applications (2^{e} année ISBS) depuis 2004
– unité N+I EL01, bases du traitement du signal, entre 2003 et 2004

Les enseignements que j'ai effectués ou effectue sont les suivants (pour 200 heures équivalent TD) : communications numériques, bases du traitement du signal, filtrage numérique, filtrage adaptatif, compression d'images, analyse spectrale, estimation et détection, traitement d'antenne, traitement satellite et radar, théorie de l'information, compression de données

Chapitre 1

Bilan succinct

Depuis 1996 je suis enseignant-chercheur à l'ESIEE, affecté au département Signaux et Télécommunications, (avec un passage de deux années dans le département Modélisation et Simulation Numérique).

Depuis début 2008, je suis membre de l'équipe "Signal et Communications" du laboratoire d'Informatique de l'Institut Gaspard Monge (LIGM, UMR 8049, CNRS). Auparavant, je faisais partie du thème "radiocommunications numériques" de l'équipe Esycom (EA 2552).

J'ai bénéficié en 2007-08 d'une "période de perfectionnement" d'une année, possibilité prévue au statut des personnels de la CCIP. Cette période de perfectionnement s'est déroulée en partie au laboratoire des signaux et systèmes LSS, UMR 8506, (situé à Gif-sur-Yvette dans les locaux de Supelec) dans l'équipe problèmes inverses ; et en partie dans l'équipe "signal et communications" du LIGM.

Comme je le notais en préambule, les recherches que je développe comprennent deux axes très distincts, qui sont menés parfois alternativement, parfois simultanément.

1. Le traitement et la théorie de l'information, et notamment l'étude de mesures d'information ou de complexité appliqué à l'analyse, le filtrage et la restauration de signaux et images, avec des applications aux télécommunications et à la physique statistique.
2. La radiotechnique (statistique) : analyse statistique et calcul de performances pour les émetteurs-récepteurs numériques, développement et caractérisation statistique d'algorithmes de correction, calibration ou contrôle.

Les paragraphes suivants donnent un bilan succinct de ces activités, en suivant l'organisation, somme toute logique, de la suite du mémoire, qui développe plus avant ces différents thèmes. Les références bibliographiques correspondent à la liste de publications présentée page 13.

—— o 0 O 0 o ——

1.1 Traitement et théorie de l'information

Entropie et maximum d'entropie

La question de la résolution de problèmes inverses à l'aide de techniques entropiques était l'objet de mon travail de thèse. Ce travail s'est poursuivi par la suite, avec quelques développements (prise en compte de corrélation, opérateur non-linéaire). Un article de synthèse [R2] a été publié en 1999. J'ai ensuite proposé une méthode de reconstruction itérative prolongeant l'idée du maximum d'entropie sur la moyenne, et s'apparentant à un algorithme EM. J'ai accueilli une étudiante en post-doctorat sur ce sujet. Cette méthode a été présentée en [C18,C19,C21], et mériterait qu'on y revienne. La technique du maximum d'entropie sur la moyenne, et sa prolongation itérative, sont présentées dans le chapitre 2, sections 2.1 et 2.2.

Toujours concernant l'entropie de Shannon, et en vue de pouvoir traiter des données, nous nous sommes intéressés, avec C. Vignat au problème d'estimation de l'entropie. Nous avons envisagé une approche

"signal", en passant par le biais d'une technique d'estimation spectrale par modèles autorégressifs régularisés. Nous nous sommes attachés à construire un estimateur de l'entropie de Shannon, qui possède des caractéristiques récursives, ce qui permet de l'employer « en ligne ». Nous avons envisagé quelques applications (séparation de sources, déconvolution, détection de ruptures) et décrit l'apport d'un tel estimateur pour ces applications. [R3,C14]. Un autre estimateur, reposant sur une manipulation des histogrammes, a été présenté en [C15].

Comme il est parfois difficile de manipuler l'entropie de Shannon, nous nous sommes penchés sur l'entropie de Rényi comme alternative pour une utilisation pratique en traitement de signal. Nous avons pu donner une inégalité sur la transformation de l'entropie de Rényi par un système linéaire et proposer une application de cette inégalité sur un problème de déconvolution MIMO. Ceci est présenté dans la communication [C17] et décrit dans la section 2.5.

Une application des mesures d'information pour la caractérisation de séquences d'ADN, et plus exactement la localisation de séquences de promoteurs de transcription a été présentée dans [C24]. Nous y avons étudié notamment un classificateur de Bayes naïf qui permet de détecter les séquences promotrices. Ce classificateur peut être relié à des tests sur des divergences de Kullback. Une suite indirecte de ce travail a été ma contribution à [R7], où j'ai effectué différents calculs statistiques qui ont permis de corroborer le modèle présenté.

Entropies de Rényi-Tsallis

Je me suis ensuite intéressé de manière plus approfondie aux entropies de Rényi, de Tsallis, à leurs interactions, aux justifications que l'on peut leur trouver et aux propriétés des lois à maximum d'entropie de Rényi-Tsallis. Ces travaux sont présentés dans le chapitre 3. Resituons d'abord en quelques mots le contexte.

Pour deux distributions univariées de densités P et Q, la divergence de Rényi (1961) est

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log \int_{\mathcal{D}} P(x)^\alpha Q(x)^{1-\alpha} dx, \tag{1.1}$$

où α est un réel positif, et $\mathcal{D}$ le domaine de définition de l'intégrale. Par la règle L'Hospital, la divergence de Kullback est obtenue pour $\alpha \to 1$.

Les applications et domaines d'intérêt pour l'entropie de Rényi sont nombreux : communication et codage, data mining, détection, segmentation, classification, test d'hypothèses, alignement d'images, etc. L'entropie de Rényi joue aussi un rôle central dans la théorie des multifractales. En physique statistique, suivant un article fondateur de Tsallis (1988), une communauté en statistiques nonextensives ("nonextensive statistics") s'est développée pour étudier une entropie comparable, et les distributions à maximum d'entropie associées, dites distributions de Tsallis

$$P_\nu(x) = \frac{\left[1 + \gamma\,(x - \bar{x})\right]^\nu}{Z_\nu(\gamma, \bar{x})}, \tag{1.2}$$

avec $Z_\nu(\gamma, \overline{x})$ la fonction de partition. Ces distributions, qui exhibent un comportement en loi puissance coïncident remarquablement avec certains jeux de données expérimentales, et semblent apparaître dans des situations à la limite de l'équilibre, ou avec un équilibre déplacé. Elles apparaissent également dans d'autres domaines : réseaux, fiabilité, climatologie, imagerie radar, économétrie...

Dans [C23], j'ai montré que les distributions de Tsallis (éventuellement généralisées avec une mesure de référence $Q(x)$) peuvent être déduites à partir d'une approche classique de maximum d'entropie, en prenant en compte une contrainte figurant un système dont l'équilibre est déplacé. Une version dérivée de cette approche a également été présentée dans une revue de physique [R11]. Le résultat me semble avoir une certaine importance car il permet de relier les distributions de Tsallis au maximum d'entropie habituel, en introduisant une contrainte qui peut d'ailleurs avoir un sens en théorie du codage. Ceci est présenté et argumenté plus avant dans la section 3.2.

Il est également intéressant de noter que dans ce contexte, la divergence de Rényi apparaît en sous-produit et la nouvelle solution d'équilibre est solution du problème

$$\mathcal{F}_\alpha^{(C \text{ resp. } G)}(m) = \left\{ \begin{array}{r} \min_P D_\alpha(P||Q) \\ \text{s.c. } (C) \;\; m = E_P[X] \\ \text{ou } (G) \;\; m = E_{P^*}[X] \\ \text{et} \quad \int_{\mathcal{D}} P(x)dx = 1 \end{array} \right. \tag{1.3}$$

On a considéré une contrainte de moyenne classique (C) et une contrainte "généralisée" (G) telle que celle qui est employée par les physiciens de la communauté de physique statistique nonextensive. À l'optimum, la solution du problème est une fonctionnelle de m, $\mathcal{F}_\alpha^{(C)}(m)$ ou $\mathcal{F}_\alpha^{(G)}(m)$.

Il est alors légitime de s'intéresser aux propriétés et caractérisations de ces "fonctionnelles entropiques". Ceci a fait l'objet du travail présenté en [R9]. J'y ai étudié et donné les formes et propriétés des solutions à maximum d'entropie, et des deux fonctionnelles associées, indexées par le paramètre α. La forme exacte (1.2) des distributions solution a été donnée et discutée. On a montré que les fonctionnelles $\mathcal{F}_\alpha^{(.)}(m)$ sont simplement liées aux fonctions de partition. Les propriétés générales de ces fonctionnelles, comprenant la non-négativité, la convexité, ont été établies. On a aussi montré comment les solutions peuvent être obtenues numériquement, quand bien même la formulation fait intervenir des fonctions implicites. Ces fonctionnelles entropiques peuvent servir de fonction objectif à optimiser (où l'on peut utiliser le logiciel [V1]), ou de potentiels dans différentes applications. Une divergence dans l'espace objet, qui se réduit à une divergence de Bregman quand $\alpha \to 1$ a été introduite. Différents exemples ont été traités, et on a vérifié que lorsque $\alpha \to 1$, on retrouve des entropies classiques. Ce travail est présenté dans la section 3.4, et l'article est annexé page 108 et suivantes.

La distribution de Tsallis (1.2) peut être aussi obtenue comme un mélange de distributions en faisant varier, suivant une loi gamma, le paramètre naturel d'une loi exponentielle. À partir de cette remarque, une théorie des 'Superstatistiques' a fait l'objet d'une série de publications depuis 2003. Pour expliquer comment le paramètre naturel d'une loi de Boltzmann-Gibbs pouvait fluctuer, j'ai proposé de décrire l'équilibre comme une solution à Maximum d'entropie avec contraintes aléatoires. Dans ce cadre, les fluctuations des contraintes entraînent d'autres fluctuations sur le paramètre de Lagrange associé, ce qui ramène aux 'superstatistiques'. J'ai ensuite considéré l'exemple naturel dans lequel les fluctuations sur les contraintes se font suivant une loi Gamma. Dans ce cas, la distribution résultante est une K-distribution, dont il se trouve qu'elle a des propriétés intéressantes et de nombreuses applications en imagerie radar, ultrasons, scintillation optique, etc. Ce travail a été décrit dans [R10].

Une caractéristique importante de la distribution de Tsallis (1.2) est son comportement en loi puissance. Compte tenu de l'importance de ces lois, souvent utilisées pour modéliser les queues de distribution, nous nous sommes intéressé aux relations entre une distribution et sa queue. Plus exactement, nous avons recherché quelles seraient les queues à maximum d'entropie, sous une certaine contrainte de distance à la distribution initiale. Le lecteur ne sera pas surpris par le fait que ces queues à maximum d'entropie sont justement les distributions de Tsallis. Ceci est est décrit dans [R11].

En réalité, ceci n'est pas complètement fortuit : avec Christophe Vignat, nous avons observé qu'un théorème sur la distribution des excès, le théorème de Pickands (1975), montre que (dès que les distributions sont dans le domaine d'attraction des distributions extrêmes) alors la distribution des excès converge vers la distribution de Pareto Généralisée, qui n'est autre que la distribution de Tsallis. Dans [C30], nous montrons que la distribution des excès converge vers une distribution de Tsallis, et que dans le même temps, la solution d'un problème de maximum d'entropie généralisée converge vers cette même solution, ce qui justifie asymptotiquement l'utilisation des entropies de Rényi-Tsallis.

Ceci fournit en outre une explication pour l'apparition des distribution de Tsallis dans un certain nombre de domaines [R12] : la distribution des excès converge vers une distribution de Tsallis et cette distribution est de plus stable par seuillage. Dès qu'un système comprend un seuillage, ou une série de seuillages d'une distribution sous-jacente, alors la loi limite converge vers une distribution de Tsallis, et cette loi d'équilibre pourra être décrite comme une distribution à maximum d'entropie de Rényi-Tsallis.

La poursuite de ces travaux comprend le prolongement de ces résultats au cas multivariable, avec en par-

ticulier l'étude des dépendances implicites qui apparaissent dans les lois à maximum d'entropie de Rényi. Ceci pourrait rejoindre les problématiques de "longues dépendances". À partir de cette idée d'analyse de dépendances statistiques dans les lois, j'ai encadré (avec Ali Djafari) un étudiant de Master Recherche sur le thème "apport de la théorie des copulas pour la reconstruction en tomographie" (les copulas permettant de mesurer des dépendances statistiques). Deux communications [C33, C36] ont été présentées à l'issue de ce travail.

L'existence d'une géométrie particulière sous jacente à la famille (1.2), à l'image de la géométrie des familles exponentielles, est également un sujet d'intérêt. La divergence de Bregman déformée que l'on a obtenue dans [R9] est un premier élément. La communication *invitée* [C29] présente une synthèse de quelques idées sur des raisons possibles de l'apparition des entropies de Rényi-Tsallis et des distributions associées. D'autres perspectives seront évoquées dans le corps du texte et dans le chapitre 7.

Information de Fisher

Plusieurs contributions concernent une autre mesure d'information, historiquement la première : l'information de Fisher. Ces contributions sont données dans le chapitre 4 et les articles en rapport [R4,R5,RS1] sont annexés. L'information de Fisher est d'une importance fondamentale en théorie de l'estimation. Elle sert également dans des problèmes d'inférence et d'interprétation de nombreux processus physique. Avec Christophe Vignat, nous avons considéré les relations entre l'information de Fisher et l'entropie de Shannon et défini le fameux "Fisher-Shannon information plane", qui a fait école. Ce plan d'information permet de caractériser les distributions de signaux, par exemple, à partir de l'analyse conjointe de leurs informations de Fisher et de Shannon. Ceci est discuté dans l'article [R4].

Nous nous sommes également intéressés à la transformation de l'information de Fisher par des systèmes singuliers. Plus spécifiquement, nous avons étudié l'extension au cas multivarié, et aux transformations non inversibles, de l'inégalité sur l'information de Fisher. Nous avons alors pu retrouver, avec des démonstrations simplifiées, les principaux résultats de Zamir sur le sujet. Ces contributions ont été publiées dans [R5].

Plus récemment, nous nous sommes intéressés à la minimisation de l'information de Fisher sous contrainte de variance et de support. Ceci revient à rechercher et caractériser les distributions de support limité, dont l'information de Fisher (relative au paramètre de localisation) est minimale, ce qui indique que le paramètre correspondant, par l'inégalité de Cramér-Rao, est le plus délicat à estimer. L'intérêt est bien sûr que dans nombre de cas pratiques, on sait que les distributions sont à support limité, par exemple à un intervalle, ou aux réels positifs. Dans l'article [RS1], actuellement en révision mineure, nous présentons les expressions explicites des solutions sur $\mathbb{R}^+$ et sur un intervalle, décrivons leur comportement, et celui de l'information de Fisher associée.

1.2 Radiotechnique

Je participe depuis quelques années à différents travaux autour des architectures radio. Mes intérêts et apports dans ces recherches sont en particulier les aspects modélisation statistique, et le développement de solutions issues des techniques de traitement du signal pour la correction et la synchronisation.

Différents problèmes d'architecture

Une des possibilités est d'aller vers des architectures entièrement numériques. Dans cette direction, j'ai coencadré, avec G. Baudoin (directrice de thèse), la thèse Cifre de C. Joubert en convention Cifre avec ST Microelectronics, (RF Expertise Center - Cellular Terminal Division), sur l'analyse et l'optimisation d'architectures fondées sur une boucle à verrouillage de phase numérique. Nous avons en particulier proposé un modèle comportemental complet du système, qui permet de le simuler avec une charge en calcul très raisonnable. L'article de congrès [C22] présente ce modèle. Nous avons également développé des algorithmes rapides permettant d'alléger l'implantation [C26], et calculé explicitement la dsp du

bruit de phase correspondant à la structure [C27]. Un prototype d'un émetteur FM employant cette PLL et implantant sur silicium les algorithmes développés a été réalisé et validé. Quelques éléments sur ces différentes contributions sont données dans la section 5.2 du chapitre 5.

Les contraintes posées pour la réalisation des émetteurs radio amènent à sélectionner des solutions de type polaire, où le signal est décomposé en un signal d'enveloppe et un signal de phase (le cosinus de la phase). Cette décomposition polaire, plutôt que la classique décomposition en phase et quadrature, permet d'utiliser des amplificateurs à haut rendement, puisque le signal présenté à l'amplificateur est de type enveloppe constante (l'enveloppe est restaurée en sortie en modulant la tension d'alimentation). Avec Corinne Berland, nous avons proposé d'utiliser un codage sigma-delta qui permet d'obtenir des performances comparables à l'aide de dispositifs existants. Cette modification de l'architecture EER a abouti à une nouvelle architecture d'émetteur complètement numérique. (C. Berland a encadré une étudiante en convention Cifre avec ST Microelectronics Crolles pour la réalisation matérielle de cette architecture dans le cadre d'un financement CIFRE). Un brevet (en une version nationale 2005 et une internationale 2006) a été déposé [V2]. Dans le cadre de ce travail, le but étant de retrouver un signal modulé « normal », la restauration de la variation de l'enveloppe se fait en sortie de l'amplificateur de puissance par une simple opération de filtrage. L'avantage de cette architecture réside dans le fait que le signal à l'entrée de l'amplificateur est à enveloppe constante. L'amplification peut être réalisée par tout type d'amplificateur (A, B , AB, C, E) et permet d'optimiser les rendements. Ceci est décrit dans l'article [R6] qui est présenté en annexe page 166 et suivantes, ainsi que dans la section 5.1 du chapitre 5.

Du point de vue des émetteurs, il convient de contrôler les différentes dérives possibles (température, alimentation, vieillissement) et stabiliser le fonctionnement, au cours du cycle de vie du mobile, et ce, d'autant plus que les fréquences de fonctionnement sont élevées. Il convient également de concevoir des solutions qui prennent en compte la variabilité statistique des composants et réalisations, et comprennent une dimension fiabilité. Des solutions sont étudiées pour un oscillateur à très hautes fréquences réalisé chez Nxp par Pierre Guillot dans le cadre de sa thèse, à l'encadrement de laquelle je participe (Coencadrement C. Berland/J.-F. Bercher). Ces aspects sont présentés ici en 5.3.

Une première étude sur la réalisation d'une référence très haute fréquence – 2GHz, construite autour d'un oscillateur BAW (*Bulk Acoustic Wave*), est présentée dans [C31], où l'on décrit les caractéristiques du système, notamment le dimensionnement d'une banque de capacités qui servira à contrôler la fréquence de l'oscillateur. La réalisation correspondante est décrite dans [C34] et [R14], où l'on présentera également des résultats de mesures. La caractérisation du système impose d'estimer ses différents paramètres, et de poursuivre leur évolution éventuelle. Ensuite, il s'agit de contrôler le système via un modèle de son évolution. Pour ces étapes d'estimation et de contrôle, nous développons une solution fondée sur une modélisation d'état et un filtrage de Kalman.

Du point de vue des récepteurs, dans le front-end avant le filtre de Nyquist, c'est-à-dire en amont des algorithmes de traitement bien connus, différents problèmes d'optimisation peuvent être abordés. Il est possible technologiquement d'échantillonner très tôt le signal RF et de disposer de capacités de traitement numérique du signal. Les points qui sont d'intérêt sont : l'échantillonnage RF, avec ou sans parallélisme, l'optimisation de la quantification, la suppression d'interférences, etc. Une thèse Cifre avec Nxp a débuté en octobre dernier sur ce thème (thèse d'Amandine Lesellier - coencadrement J.-F. Bercher/O. Venard), avec une application aux récepteurs de télévision numérique multicanaux. Le problème est de pouvoir accéder simultanément à plusieurs canaux, 5 typiquement, situés quelque part dans une bande de 1 GHz. Le problème est donc de trouver une procédure d'échantillonnage très large bande, permettant de préserver plusieurs canaux de localisations aléatoires. Pour ce travail, on a commencé à étudier des solutions autour de l'utilisation de bancs de filtres hybrides analogique/numérique, avec une condition de reconstruction parfaite. On s'intéresse également aux possibilités d'utiliser des résultats sur l'échantillonnage non-uniforme, ou aléatoire, qui permet de dépasser la condition d'échantillonnage de Shannon (Whittaker).

Synchronisation des voies dans les architectures polaires

Dans l'architecture polaire que l'on a évoquée plus haut, le signal d'enveloppe et le signal de phase subissant des traitements différents, une désynchronisation entre l'enveloppe et la phase peut apparaître. Cette désynchronisation entraîne une remontée du plancher de bruit et une augmentation de l'EVM (Error Vector Magnitude). Ce problème devient d'autant plus crucial que la période symbole diminue, avec l'augmentation des débits. Par ailleurs, les caractéristiques analogiques sont affectées par des dérives dans le temps, par une variation avec la température et par un effet de vieillissement. Enfin, les solutions sont de plus en plus multimodes et les caractéristiques peuvent varier lorsque l'on passe d'un mode à un autre.

Avec Corinne Berland, nous avons travaillé sur les possibilités de compenser ces retards entre voies. Ceci est l'objet du chapitre 6, à partir de la page 53.

Nous avons proposé un premier algorithme de correction dans [C25]. En nous appuyant sur les statistiques de signaux, nous avons montré d'une part que la synchronisation ne peut pas être effectuée en aveugle et nécessite une boucle de retour. Nous avons également discuté d'une procédure d'identification dans [C28]. Le principe de correction est simplement d'ajuster des paramètres afin de minimiser l'erreur quadratique d'entrée sortie, à l'aide d'un algorithme stochastique. La procédure emploie une étape d'interpolation numérique qui a été discutée. Nous avons caractérisé analytiquement le comportement de l'algorithme (vitesse de convergence, biais, variance) à l'aide d'un 'toy model', et ces résultats ont été confrontés à des simulations numériques. L'ensemble a été validé sur un modèle de modulation OFDM [R8].

Nous avons ensuite développé une version sous optimale, qui permet d'alléger notablement l'implantation numérique comme la réalisation en hardware, en évitant de devoir effectuer une démodulation en quadrature qui peut être remplacée par une conversion directe. Le domaine de convergence de ce nouvel algorithme a été caractérisé. Nous avons également vérifié analytiquement que ces solutions sont robustes à une mauvaise connaissance du gain ou à un offset de phase ; mais bien entendu au prix d'une augmentation de la variance d'estimation. Nous avons alors développé une extension de la procédure permettant de prendre en compte et compenser ces méconnaissances. Enfin, nous avons étudié le comportement de cette solution en contexte non stationnaire. Un article reprenant ces différents résultats est actuellement soumis [RS2]. Les deux publications [R8], et [RS2] sont annexées.

Nous avons d'ores et déjà prolongé ces résultats, analytiquement et en simulation, à l'architecture LINC, potentiellement très performante, mais qui est confrontée aux mêmes types de désynchronisation, peut-être de manière plus aigüe encore. Une autre perspective consistera a également prendre en charge des ondulations résiduelles du gain dans la bande passante.

La conception et l'optimisation des émetteurs numériques semble ainsi pouvoir bénéficier de l'apport d'outils de traitement du signal pour l'optimisation de la "dirty RF" : c'est le point de vue présenté dans [C32].

1.3 Publications

Articles de revue avec comité de lecture

Articles soumis :

[RS1] J.-F. Bercher and C. Berland, "Adaptive delays alignment in polar transmitter architecture," *IEEE trans. on Circuits and Systems I*, submitted, April 2009.

Articles publiés :

[R15] J.-F. Bercher and C. Vignat, "On minimum Fisher information distributions with restricted support and fixed variance," *Information Sciences*, vol. 179,pp. 3832-3842, Nov. 2009.

[R14] P. Guillot, P. Philippe, C. Berland, J.-F. Bercher and P. Gamand, "Low Noise High Resolution Silicon-Based High Frequency Oscillator", *Electronics Letters*, vol. 45, 2009, pp. 914-916, august 2009.

[R13] J.-F. Bercher, "Source Coding with Escort Distributions and Rényi Entropy Bounds," *Physics Letters A*, vol. 373, pp. 3235-3238, august 2009.

[R12] J.-F. Bercher and C. Vignat, "A new look at q-exponential distributions via excess statistics", *Physica A*, vol. 387, n°22, pp 5422-5432, september 2008.

[R11] J.-F. Bercher, "Tsallis distribution as a standard maximum entropy solution with 'tail' constraint", *Physics Letters A*, vol. 372, n°35, pp 657-5659, august 2008.

[R10] J.-F. Bercher, "Maximum entropy with fluctuating constraints : The example of K-distributions", *Physics Letters A*, vol. 372, n°24, pp 4361-4363, june 2008.

[R9] J.-F. Bercher, "On some entropy functionals derived from Rényi information divergence", *Information Sciences*, vol. 178, pp. 2489-2506, june 2008.

[R8] J.-F. Bercher and C. Berland, "Envelope and phase delays correction in an EER radio architecture", *Analog Integrated Circuits and Signal Processing*, vol. 55, pp. 21-35, Apr. 2008.

[R7] B. Duriez, P. Duquesnoy, E. Escudier, A.-M. Bridoux, D. Escalier, I. Rayet, E. Marcos, A.-M. Vojtek, J.-F. Bercher and S. Amselem, "A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, n°9, pp. 3336-3341, february 2007 ; and "Corrections", vol. 104, n°15, p. 6490, april 2007.

[R6] C. Berland, I. Hibon, J.-F. Bercher, M. Villegas, D. Belot, D. Pache, V. Le Goascoz "New transmitter architecture for nonconstant envelope modulation", *IEEE Trans. on Circuits and Systems*, vol. 53, Issue 1, pp 13 - 17, january 2006.

[R5] C. Vignat, J.-F. Bercher,"On Fisher information inequalities and score functions in non-invertible linear systems", *JIPAM. Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 4, Article 71, 9 pages, november 2003.

[R4] C. Vignat, J.-F. Bercher,"Analysis of signals in the Fisher–Shannon information plane", *Physics Letters A*, vol. 312, pp. 27–33, june 2003.

[R3] J.-F. Bercher, C. Vignat, "Estimating the entropy of a signal with applications", *IEEE Trans. on Signal Processing*, vol. 48, n° 6, pp. 1687-1694, june 2000.

[R2] G. Le Besnerais, J.-F. Bercher, G. Demoment, "A new look at the entropy for solving linear inverse problems", *IEEE Trans. on Information Theory*, vol. 45, n°5, pp.1565-1578, july 1999.

[R1] J.-F. Bercher et P. Larzabal, « Traitement d'antenne avec adéquation de modèle, » Traitement du Signal, numéro spécial *Calibration d'antenne*, volume 10, n° 5-6, pp. 349–359, 1995.

Colloques avec comité de lecture et actes

[C36] D.-B. Pougaza, A. Mohammad-Djafari, J.-F. Bercher, « Utilisation de la notion de copule en tomographie », Actes du colloque Gretsi, Dijon, septembre 2009.

[C35] C. Vignat and J.-F. Bercher, "Geometric aspects of the non-extensive statistical theory", MaxEnt 2009, July 2009, Oxford, Mississippi.

[C34] P. Guillot, P. Philippe, C. Berland, and J.-F. Bercher, "Faisabilité de référence haute fréquence pour les architectures RF," in *Journées Nationales Microondes 2009*, may 2009.

[C33] A. Mohammad-Djafari, D. Pougaza, and J.-F. Bercher, "Copula and Tomography," International Conference on Computer Vision Theory and Applications, Lisbon, Portugal : 2009.

[C32] C. Berland, J.-F. Bercher, O. Venard, "Digital Signal Processing techniques to compensate for RF imperfections in Advanced Transmitter Architectures", European Wireless Technology Conference (EuWit2008), Amsterdam, october 2008.

[C31] P. Guillot, C. Berland, J.-F. Bercher, P. Philippe, "A 2GHz 65nm CMOS digitally-tuned BAW oscillator", 15th IEEE International Conference on Electronics (ICECS2008), Circuits and Systems, Malta, september 2008.

[C30] J.-F. Bercher and C. Vignat, "An entropic view of Pickands' theorem", IEEE International Symposium on Information Theory (ISIT2008), Toronto, july 2008

[C29 J.-F. Bercher, "Possible rationales for Rényi-Tsallis entropy maximization", International Workshop on Applied Probabilities (IWAP2008), Compiègne, july 2008.

[C28] C. Berland and J.-F. Bercher, "Adaptive time mismatches identification and correction in polar transmitter architecture", European MicroWave 2007, Munich, 8-12 october 2007.

[C27] G. Baudoin, J.-F. Bercher and C. Joubert, « Calcul théorique du bruit de phase dû au dithering dans un oscillateur contrôlé numériquement », Journées Nationales Microondes, Toulouse, 23-25 Mai 2007

[C26] C. Joubert, J.-F. Bercher, G. Baudoin, T. Divel, S. Ramet, P. Level, "Contributions to the analysis and design of an ADPLL", Proceedings of the 13th IEEE International Conference on Electronics, Circuits, and Systems, ICECS2006, pp. 322-325, december 2006.

[C25] J.-F. Bercher and C. Berland, "Envelope/phase delays correction in an EER radio architecture", Proceedings of the 13th IEEE International Conference on Electronics, Circuits, and Systems, ICECS2006, pp. 443-446, december 2006.

[C24] J.-F. Bercher, P. Jardin et B. Duriez, "Bayesian classification and entropy for promoter prediction in human DNA sequences", AIP Conference Proceedings, vol. 872, pp. 235-242, Twenty sixth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 2006

[C23] J.-F. Bercher, " An amended MaxEnt formulation for deriving Tsallis factors, and associated issues", AIP Conference Proceedings, vol. 872, pp. 441-448, Twenty sixth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 2006, ArXiv.org

[C22] C. Joubert, J.-F. Bercher, G. Baudoin, T. Divel, S. Ramet, P. Level, "Time Behavioral model for Phase Domain ADPLL based frequency synthetizer" IEEE Radio and Wireless Symposium (RWS 2006), january 2006.

[C21] P. Holejsovska, Z. Peroutka, J. Cengery, J.-F. Bercher, "Continuous Non-Invasive Blood Pressure Measuring Based on the Reconstruction by Maximum Entropy on the Mean Method" In Jirí Jan, Jirí-Kozumplík, and Ivo Provazník, editors, BIOSIGNAL 2004 : 17-th Biennial International Eurasip Conference Proceedings, pages 121-123, Brno, Czech Republic, June 2004.

[C20] J.-F. Bercher, A. Diet, C. Berland,G. Baudoin, M. Villegas, "Monte carlo Estimation of Time mismatch Effect in an OFDM EER Architecture", IEEE Radio & Wireless Conference RAWCON 2004, sept 2004.

[C19] P. Holejšovskà, Z. Peroutka, J.-F. Bercher, "Adaptive signal processing using maximum entropy on the mean method and Monte Carlo analysis" In : MATLAB 2003, pp 191-196, ISBN 80-7080-526-9.

[C18] P. Holejšovskà, J.-F. Bercher, "Elimination of the Measurement System Impact on the Measured Data by Maximum Entropy on the Mean Method", Diagnostika '03, Plzen, pp. 192-195, september 2003, ISSBN 80-7082-952-4.

[C17] J.-F. Bercher, C. Vignat, "Renyi entropy convolution inequality with Applications", Eusipco 2002, Toulouse, France, September 2002.

[C16] C. Vignat, J.-F. Bercher, "Matrix Fisher Information Inequalities for Non-Invertible Linear Systems", IEEE International Symposium on Information Theory, Lausanne, Switzerland, June 30-July 5, 2002.

[C15] C. Vignat, J.-F. Bercher, "Un estimateur récursif de l'entropie", *actes du 17e Colloque* GRETSI *sur le Traitement du Signal et ses Applications,* Vannes, septembre 1999.

[C14] J.-F. Bercher, C. Vignat, "Estimating the entropy of a signal with applications", Proceedings. ICASSP99 (Cat. No.99CH36258), 1999, Page 1705-1708 Volume 3.

[C13] C. Leal, C. Meihlac, A. Pesme, J.-F. Bercher, C. Vignat, "Recovering binary data transmitted over unknown communication channels", *in Proc. of DSPCONF Texas-Instrument*, Noisy-le-Grand, sept. 1998, pp. 119-124.

[C12] D. Janu, G. Baudoin, J.-F. Bercher, O. Venard, "Design of a CDMA simulator and implementa tion on a TMS320601", *in Proc. of DSPCONF Texas-Instrument*, Noisy-le-Grand, Sept. 1998, pp. 101-106.

[C11] C. Heinrich, J.-F. Bercher et G. Demoment, « The Maximum Entropy on the Mean Method, Correlations, and Implementation Issues,» 16th Conference on Maximum Entropy and Bayesian Methods, South Africa, août 1996, 10 pages.

[C10] J.-F. Bercher et C. Heinrich, « Construction de mesures de divergence. Application à la résolution de problèmes inverses linéaires, » 15e Colloque GRETSI sur le Traitement du Signal et ses Applications, Juan-les-Pins, septembre 1995, 4 pages, et une version longue.

[C9] C. Heinrich, J.-F. Bercher, G. Le Besnerais et G. Demoment, « Méthode du maximum d'entropie sur la moyenne et mélanges de distributions, » 15e Colloque GRETSI sur le Traitement du Signal et ses Applications, Juan-les-Pins, septembre 1995, 4 pages.

[C8] C. Heinrich, J.-F. Bercher, G. Le Besnerais et G. Demoment, "Restoration of spiky signals : a new optimal estimate and a comparison," IEEE International Conference on Acoustic Speech and Signal Processing, mai 1995, 4 pages.

[C7] J.-F. Bercher, G. Le Besnerais et G. Demoment, « Building Convex Criteria for solving linear inverse problems," Inverse Problems and Applications, HoChiMinh City, Vietnam, 17-19 janvier 1995, 12 pages.

[C6] J.-F. Bercher, G. Le Besnerais et G. Demoment, « The Maximum Entropy on the Mean Method, Noise, and Sensitivity," 14th Conference on Maximum Entropy and Bayesian Methods, Cambridge, 1-5 août 1994. Publié dans *Maximum entropy and Bayesian Methods*, Kluwer Academic Publishers, The Nederlands, 1996, pp. 223-232.

[C5] G. Le Besnerais, J.-F. Bercher et G. Demoment, « Probabilistic Issues in Fourier Synthesis," First School on Scientific Computation in Systems, Jouy-en-Josas, 6-8 décembre 1993, 24 pages.

[C4] J.-F. Bercher, « Un algorithme de factorisation bispectrale, » 14e Colloque GRETSI sur le Traitement du signal et ses Applications, Juan-les-Pins, septembre 1993, pp 117–120.

[C3] B. Lumeau, J.-C. Pesquet, J.-F. Bercher, L. Louveau, "Optimization of Bias-variance Trade-Off in Non parametric Spectral Analysis by decomposion into Wavelets Packets," International Conference Wavelets and Applications, Toulouse, 8-12 juin 1992, 6 pages. Publié dans "Progress in Wavelet Analysis and Applications", Y. Meyer and S. Roques ed., Éditions Frontières, 1993.

[C2] B. Lumeau et J.-F. Bercher, « Extension de la méthode de Pisarenko et ses dérivées au cas de signaux vectoriels et applications à la localisation de sources, » 13e Colloque GRETSI sur le Traitement du Signal et ses Applications, Juan-les-Pins, 16-20 septembre 1991, pp 361–364.

[C1] B. Lumeau et J.-F. Bercher, « Critère optimisant un compromis biais-variance pour estimer l'ordre d'un processus autorégressif mono ou multivariable, » 13e Colloque GRETSI sur le Traitement du Signal et ses Applications, Juan-les-Pins, 16-20 septembre 1991, pp 321–324.

Autres publications

[1] P. Holejovskà, J.-F. Bercher. "Adaptive signal processing : maximum entropy on the mean method" In : Elektrotechnika a informatika 2003, pp. 35-38, ISBN 80-7082-992-3

Valorisation

[V2] C. Berland, J. F. Bercher, I. Hibon, M. Villegas, D. Belot, D. Pache, V. Le Goascoz, "Digital Transmitter Architecture", Brevet FR0405636, publié le 25/05/2004 ; Brevet international US2006034391A1, publié le 16/02/2006.

[V1] J.-F. Bercher, S. Brette, H. Carfantan, J.-F. Giovannelli, C. Heinrich, J. Idier, T. Martin, C. Soussen, « Gradient à pas Adaptatif avec Corrections », logiciel déposé par le CNRS auprès de l'Agence de Protection des Programmes sous le N°00258-01, juin 2005.

Thèse

[T1] J.-F. Bercher, « Développement de critères de nature entropique pour la résolution de problèmes inverses linéaires, » Thèse de doctorat, Université Paris-Sud, 1995.

Livres

Participation à l'ouvrage collectif d'enseignement (à paraître en septembre 2009) :

[L2] « *Antennes* », sous la direction d'O. Picon, Dunod 2009. Première édition, 2009.

Auteur du chapitre traitement d'antenne.

Participation à l'ouvrage collectif d'enseignement :

[L1] « *Systèmes de communications numériques – Modélisation et simulation* », sous la direction de G. Baudoin, Dunod 2002. Seconde édition, 2007.

Auteur ou coauteur des chapitres :

- Récepteur optimal,
- Égalisation,
- Synchronisation (avec J.-M. Brossier).

Colloques sans actes, journées scientifiques

- J.-F. Bercher, « Codage de source pour des mesures de longueur moyenne généralisées avec pour bornes minimales les entropies de Rényi-Tsallis » Rencontre franco-suisse sur le bruit et les non linéarités, Grenoble, 10 Juin 2009.
- J.-F. Bercher, « Fonctionnelles entropiques associées à la divergence de Rényi », Séminaire image, Centre Mathématique de Jussieu-Chevaleret, 4 novembre 2008.

– J.-F. Bercher et C. Vignat, « Levy distributions and the maximization of Rényi-Tsallis entropy", in "Power laws in probabilities and statistics", CIRM, march 22-26, 2004.
– J.-F. Bercher et C. Vignat, « Estimation de l'entropie d un signal et applications, » GdR Information, signal et images (GT 1 : Méthodes Entropiques en Signal et Image), 10 décembre 1998 ; et sur le même thème (GT2 Séparation de sources) 29 avril 1999.
– J.-F. Bercher, « Critères de type entropique, » Journée thématique du GRCE-AFCET (Groupe de Reconnaissance en Communication Ecrite), 06 février 1996.
– J.-F. Bercher, « Critères entropiques et problèmes inverses, » Journée scientifique « Information & signal » du GdR Information, signal et images, 10 novembre 1995.
– J.-F. Bercher, « Critères entropiques et optimisation convexe pour la résolution de problèmes inverses linéaires, » journée scientifique « optimisation », École Nationale Supérieure des Télécommunications de Bretagne, 18 mai 1995.
– J.-F. Bercher, C. Heinrich, G. Le Besnerais et G. Demoment, « Utilisation de la méthode du maximum d'entropie sur la moyenne en traitement du signal et de l'image, » journée maximum d'entropie sur la moyenne et grandes déviations, Université d'Évry, 1er avril 1995.
– J.-F. Bercher et C. Heinrich, « Entropies et reconstruction d'images, » 3^{e} Colloque jeunes chercheurs en physique, Campus d'Orsay, 27-28 mars 1995.

Chapitre 2

Entropie et maximum d'entropie

L'entropie et le principe du maximum d'entropie sont parmi mes sujets de réflexion majeurs depuis mon travail de thèse. Le rôle central de l'entropie de Shannon en théorie de la communication est bien connu, tout comme le rôle de l'entropie de Boltzmann-Gibbs en physique statistique et en thermodynamique. Dans deux articles fondamentaux [3, 4], Jaynes discute de l'interaction entre la théorie de l'information et la mécanique statistique. En particulier, il introduit le principe du maximum d'entropie : *Information theory provides a constructive criterion for setting up probability distributions on the basis of partial knowledge, and leads to a type of statistical inference which is called the maximum entropy estimate. It is the least biased estimate possible on the given information ; i.e., it is maximally noncommittal with regard to missing information*. Dans ce chapitre, on discute de méthodes d'inférence s'appuyant sur le maximum d'entropie, de caractérisations du principe lui même, puis finalement de l'intervention de l'entropie comme mesure de la complexité d'une distribution permettant le traitement de données. Le premier paragraphe, 2.1 reprend le travail [5] sur le maximum d'entropie sur la moyenne débuté lors de la préparation de ma thèse de doctorat. Le second paragraphe 2.2 présente une technique de maximum d'entropie sur la moyenne itératif [6, 7]. Dans le paragraphe 2.3 on rapporte quelques idées sur le cas du maximum d'entropie avec des contraintes aléatoires, et l'apparition de K-distributions [8]. Dans le paragraphe suivant, section 2.4, on présente un estimateur de l'entropie de Shannon et des exemples d'application [9]. Enfin, dans le dernier paragraphe, on présente une inégalité sur l'entropie de Rényi que l'on utilise dans le cadre de problèmes de déconvolution [10].

2.1 Maximum d'entropie sur la moyenne

Le travail sur le maximum d'entropie sur la moyenne correspond au départ à mon travail de doctorat, soutenu en 1995. Ce travail s'est poursuivi au delà, avec en particulier la parution d'un article [5] en 1999, qui se trouve un peu à la frontière d'une recherche du travail de thèse et d'une recherche autonome. Par rapport au travail de thèse, les mathématiques ont été clarifiées, et de nouveaux exemples et développements insérés (le cas d'un opérateur non linéaire, la prise en compte d'un bruit poissonnien, l'introduction de corrélations). D'autres idées et travaux ultérieurs, exposés dans la suite du document, se rattachent également à cette thématique. Je reprends donc rapidement ici le principe de la méthode et les principaux résultats obtenus, sans m'appesantir outre mesure sur les subtilités. Les lecteurs intéressés pourront consulter par exemple [5], ou l'excellent article de H. Gzyl [11]. On considère un problème inverse linéaire, comme il en apparaît dans de nombreuse situations expérimentales, qui se pose donc sous la forme simple

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}. \tag{2.1}$$

Le problème qui consiste à retrouver $\mathbf{x}$ est parfois (souvent) mal-posé (sous-déterminé, ou matrice $\mathbf{A}$ mal-conditionnée). On cherche alors à trouver une solution acceptable, par exemple en incluant des *desiderata*, ou des contraintes sur $\mathbf{x}$ (ou sur une représentation de $\mathbf{x}$) en minimisant un critère composite de la forme

$$\mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{y} - \mathbf{A}\mathbf{x}). \tag{2.2}$$

Ce type de critère peut bien entendu provenir ou être interprété comme un maximum a posteriori. Les propriétés des fonctionnelles $\mathcal{F}$ et $\mathcal{G}$ ont bien sûr de l'importance, du point de vue pratique ou pragmatique : unimodalité, convexité, support, dualité, etc. La technique du maximum d'entropie sur la moyenne permet de construire des critères convexes $\mathcal{F}(\mathbf{x})$ incorporant des caractéristiques a priori, notamment des contraintes de support et possédant une formulation duale attrayante. Considérons pour le moment le problème sans bruit, c'est-à-dire simplement le problème $\mathbf{y} = \mathbf{A}\mathbf{x}$, typiquement sous déterminé. L'idée est, comme en physique, de considérer l'objet recherché $\mathbf{x}$ comme la moyenne sous une loi. Dans la mesure où plusieurs distributions peuvent être compatibles avec les observations, on choisit de sélectionner la distribution à maximum d'entropie, dont la moyenne vérifie $\mathbf{y} = \mathbf{A}\mathrm{E}[\mathbf{X}]$. En réalité, on peut savoir que l'objet recherché possède des propriétés de concentration, ou de support ($\mathbb{R}^+$ ou un intervalle par exemple). Plutôt que de maximiser l'entropie standard, on maximisera une μ-entropie, où μ est une mesure sur le domaine concerné, ou, de manière équivalente, on cherche à minimiser la distance de Kullback-Leibler à une loi de référence, disons μ, que l'on note $D(P||\mu)$. Notons que cette démarche trouve une argumentation par le théorème de Sanov. Si P_N est la distribution empirique correspondant à un tirage de N variables aléatoires selon une loi μ, alors la probabilité Q de trouver P_N dans un ensemble $\mathcal{B}$ est, grosso-modo, et pour N grand

$$Q\left(P_N \in \mathcal{B}\right) \approx \exp\left(-N \inf_{P\in\mathcal{B}} D(P||\mu)\right). \tag{2.3}$$

On en déduit donc, en itérant le raisonnement sur des sous ensembles de $\mathcal{B}$, que la distribution absolument prépondérante dans $\mathcal{B}$ est celle qui réalise le minimum de la distance de Kullback à μ : on a concentration de toute la probabilité sur la distribution la plus proche de μ. En réalité ce qui nous intéresse c'est plutôt la probabilité de trouver une moyenne empirique x_N, c'est-à-dire la moyenne sous P_N, dans un ensemble $\mathcal{C}$. Dans ce cas, on a un résultat de grandes déviation de niveau 1, qui indique que

$$Q\left(x_N \in \mathcal{C}\right) \approx \exp\left(-N \inf_{x\in\mathcal{C}} \mathcal{F}(\mathbf{x})\right), \tag{2.4}$$

où $\mathcal{F}(\mathbf{y})$ est la fonction de taux. Ce résultat suggère maintenant de sélectionner l'élément le plus probable, celui qui réalise le minimum de $\mathcal{F}(\mathbf{x})$ sur $\mathcal{C}$. Le passage d'une problématique sur les distributions à une problématique sur les moyennes est connu comme le principe de contraction. Si on revient au problème initial, on voit que la minimisation de l'information de Kullback-Leibler sous la contrainte $\mathbf{y} = \mathbf{A}\mathrm{E}[\mathbf{X}]$ peut se traiter comme une minimisation en deux temps : d'abord une minimisation sous $\mathbf{x} = \mathrm{E}[\mathbf{X}]$, puis une minimisation sous la contrainte $\mathbf{y} = \mathbf{A}\mathbf{x}$:

$$\inf_{P:\mathbf{y}=\mathbf{A}\mathrm{E}[\mathbf{X}]} D(P||\mu) = \inf_{\mathbf{x}:\mathbf{y}=\mathbf{A}\mathbf{x}} \left[\inf_{P:\mathbf{x}=\mathrm{E}[\mathbf{X}]} D(P||\mu)\right]. \tag{2.5}$$

On voit que la minimisation de droite sur P définit en fait une fonction de $\mathbf{x}$:

$$\mathcal{F}(\mathbf{x}) = \inf_{P:\mathbf{x}=\mathrm{E}[\mathbf{X}]} D(P||\mu) \tag{2.6}$$

Il n'est pas difficile de vérifier que la fonction $\mathcal{F}(\mathbf{x})$ ainsi définie est strictement convexe (minimisation d'une fonction strictement convexe sous une contrainte linéaire). Le problème initial, sur l'espace des distributions de probabilité, devient ainsi équivalent à une minimisation sur l'espace objet

$$\inf_{P:\mathbf{y}=\mathbf{A}\mathrm{E}[\mathbf{X}]} D(P||\mu) = \inf_{\mathbf{x}:\mathbf{y}=\mathbf{A}\mathbf{x}} \mathcal{F}(\mathbf{x}) \tag{2.7}$$

Considérons encore un instant (2.6). On sait que la solution de la minimisation est dans la famille exponentielle engendrée par μ : $dP(\mathbf{x}) = \exp(\lambda^T\mathbf{x} - \log(Z(\lambda^T))d\mu(\mathbf{x})$, où $Z(\lambda^T)$ est la fonction de partition qui sert à normaliser $dP(\mathbf{x})$. Dans ce cas, et si tout va bien, on a égalité avec le problème dual associé, et

$$\mathcal{F}(\mathbf{x}) = \sup_{\lambda}\{\lambda^T\mathbf{x} - \log(Z(\lambda^T))\} = \sup_{\lambda}\{\lambda^T\mathbf{x} - \mathcal{F}^*(\lambda^T)\}. \tag{2.8}$$

La seconde égalité ci-dessus introduit la fonction $\mathcal{F}^*(\lambda^T)$. Sous cette forme, la fonction $\mathcal{F}(\mathbf{x})$ apparaît comme la conjuguée convexe, ou la transformée de Legendre-Fenchel de $\mathcal{F}^*(\lambda^T)$, et $\mathbf{x}$ et λ sont les variables conjuguées. Dans le cas présent, on identifie $\mathcal{F}^*(\lambda^T)$ au logarithme de la transformée de Laplace de la mesure μ : la transformée de Legendre-Fenchel correspondante est la transformée de Cramér de μ. On sait en outre que les variables duales sont reliées par les relations de « dualité de Legendre » : $\mathbf{x} = \mathcal{F}^{*'}(\lambda_*^T)$ et $\lambda_* = \mathcal{F}'(\mathbf{x})$. D'un autre côté, on retrouve aussi le fait que la dérivée de la log-fonction de partition, par rapport au paramètre naturel λ, n'est autre que la moyenne sous la loi.

Revenons maintenant à la résolution du problème et à l'équation (2.7). Les deux problèmes d'optimisation, dans l'espace des mesures de probabilité et dans l'espace objet, possèdent la même formulation duale :

$$\sup_{\lambda}\{\lambda^T\mathbf{y} - \mathcal{F}^*(\lambda^T\mathbf{A})\}, \tag{2.9}$$

où $\mathcal{F}^*$ est comme précédemment la log-fonction de partition de la famille exponentielle de paramètre naturel $\mathbf{A}^T\lambda$. En résolvant (2.9), on obtient ainsi à la fois la distribution à maximum d'entropie et la moyenne correspondante, qui est aussi la solution du problème de minimisation de $\mathcal{F}(\mathbf{x})$ sous contrainte, et qui rappelons le, n'est autre que la dérivée de $\mathcal{F}^*$ à l'optimum. Il est utile de remarquer que $\mathcal{F}^*$ est également une fonction convexe. On note que finalement il n'est pas nécessaire de travailler dans l'espace direct, et que la résolution se fait dans le domaine dual, généralement sans contrainte.

Associé à la mesure de référence μ, on peut ainsi trouver un critère entropique $\mathcal{F}(\mathbf{x})$, strictement convexe, qui encode certaines caractéristiques de μ. Notamment, lorsqu'on utilise une mesure à support borné, par exemple la loi uniforme ou loi de Bernoulli, ou une loi sur $\mathbb{R}_+^N$, le critère n'est défini que sur l'enveloppe convexe de la mesure, et la solution, forcément, appartient à ce domaine. Un certain nombre d'exemples de critères sont donnés et interprétés ainsi dans la publication [5], en remarquant cependant qu'il n'est pas nécessaire de disposer d'une expression explicite pour $\mathcal{F}$ puisque la résolution se fait dans le domaine dual.

Disons encore un mot sur les techniques de prise en compte du bruit d'observation : lorsqu'on relaxe la contrainte égalité $\mathbf{y} = \mathbf{A}\mathbf{x}$ à une contrainte du type $||\mathbf{y} - \mathbf{A}\mathbf{x}||^2 < c$, on peut résoudre le problème par $\inf_{\mathbf{x}} \mathcal{F}(\mathbf{x}) + \alpha||\mathbf{y} - \mathbf{A}\mathbf{x}||^2$, où α est un paramètre de Lagrange. On peut également utiliser une approche duale, dans laquelle le terme quadratique précédent se traduit simplement par l'ajout d'un terme en $\alpha||\lambda||^2$. Une autre approche consiste à définir un objet étendu $\mathbf{w} = [\mathbf{x}; \mathbf{b}]^T$, qui vérifie alors $\mathbf{y} = \mathbf{H}\mathbf{w}$ avec $\mathbf{H} = [\mathbf{A}|\mathbf{1}]$. On peut alors introduire une mesure de référence $\mu_w = \mu_x\mu_b$ sur l'objet global et utiliser la démarche du maximum d'entropie sur la moyenne. Celle-ci conduit alors à un critère composite dans l'espace objet

$$\mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{y} - \mathbf{A}\mathbf{x}), \tag{2.10}$$

où $\mathcal{F}$ et $\mathcal{G}$ sont respectivement les critères entropiques associés aux mesures μ_x et μ_b.

2.2 Maximum d'entropie sur la moyenne itératif

Il n'existe pas de connection naturelle, sinon asymptotique [12], entre la construction du maximum d'entropie sur la moyenne et une approche bayésienne. Les techniques à maximum d'entropie sont parfois, souvent même, insérées dans une démarche bayésienne en prenant comme a priori une forme $\exp \mathcal{F}(\mathbf{x})$, et la recherche d'un MAP conduit alors à la maximisation d'un critère régularisé classique. On présente ici une alternative aux techniques de prise en compte du bruit exposées dans le paragraphe précédent, avec une approche qui fait coopérer l'approche bayésienne et celle du maximum d'entropie sur la moyenne.

L'idée de base est la suivante. On a vu que l'on sait bien résoudre le problème sans bruit $\bar{\mathbf{y}} = \mathbf{A}\mathbf{x}$, pour lequel il est respectable de postuler que l'on recherche une solution comme la moyenne d'une distribution à maximum d'entropie. En réalité, l'observation est bruitée, et l'on a $\mathbf{y} = \bar{\mathbf{y}} + \mathbf{b}$.

Supposons que l'on dispose d'une estimée, disons $\bar{\mathbf{y}}_k$ de la moyenne des observations. Dans ces conditions, on peut résoudre suivant

$$\bar{\mathbf{x}}_k = \operatorname{argmin}_{\bar{\mathbf{y}}_k = \mathbf{A}\mathbf{x}} \mathcal{F}(\mathbf{x}), \tag{2.11}$$

ou de manière équivalente, le problème sur le paramètre naturel de la loi à maximum d'entropie :

$$\bar{\boldsymbol{\lambda}}_k = \text{argmax}_\lambda\{\lambda^T \bar{\mathbf{y}}_k - \mathcal{F}^*(\lambda^T \mathbf{A})\}. \tag{2.12}$$

Lorsque $\bar{\mathbf{y}}_k$ est connue, la description de notre objet $\mathbf{x}$, est ainsi donnée par la loi à maximum d'entropie, une loi exponentielle par rapport à la densité de référence $\mu(x)$, de paramètre $\bar{\boldsymbol{\lambda}}_k$. Il est alors parfaitement légitime d'utiliser cette densité, que l'on pourra noter $p(\mathbf{x}|\bar{\mathbf{y}}_k)$,

$$p(\mathbf{x}|\bar{\mathbf{y}}_k) = \exp\left(\boldsymbol{\lambda}_k^T \mathbf{A}\mathbf{x} - \log Z(\boldsymbol{\lambda}^T \mathbf{A})\right) \mu(\mathbf{x}) \tag{2.13}$$

comme densité a priori et d'appliquer la mécanique bayésienne pour aboutir à la loi a posteriori selon

$$p(\mathbf{x}|\mathbf{y}, \bar{\mathbf{y}}_k) \propto p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}|\bar{\mathbf{y}}_k). \tag{2.14}$$

À partir de cette loi a posteriori, qui a donc intégré les véritables observations $\mathbf{y}$, on peut calculer la moyenne a posteriori $\check{\mathbf{x}}_{k+1} = \mathrm{E}\left[\mathbf{X}|\mathbf{y}, \bar{\mathbf{y}}_k\right]$, et définir une nouvelle estimée de la moyenne des observations par

$$\bar{\mathbf{y}}_{k+1} = \mathbf{A}\check{\mathbf{x}}_{k+1} = \mathbf{A}\mathrm{E}\left[\mathbf{X}|\mathbf{y}, \bar{\mathbf{y}}_k\right]. \tag{2.15}$$

Il ne reste plus alors qu'à itérer le processus, tant que, par exemple, $||\bar{\boldsymbol{\lambda}}_k - \bar{\boldsymbol{\lambda}}_{k-1}||^2 > \epsilon$. On pourra alors retenir comme estimée, selon les cas, soit la moyenne a posteriori, soit la moyenne de la loi à maximum d'entropie. Comme on dispose de la loi a posteriori, il est également possible de calculer une covariance a posteriori, ou au moins des variances a posteriori.

En fait, cette procédure est analogue à un algorithme EM pour une famille exponentielle engendrée par la densité de référence initiale $\mu(\mathbf{x})$. C'est le choix de la loi a priori à maximum d'entropie (2.13) qui confère à la loi a posteriori complète cette nature de famille exponentielle (l'ajout de la vraisemblance ne modifie rien à cet aspect). Le paramètre naturel à estimer est alors $\Psi = \mathbf{A}\boldsymbol{\lambda}$, ou plus simplement λ, et dans notre cas la statistique suffisante est simplement $t(\mathbf{x}) = \mathbf{x}$. Dans ces conditions les deux étapes de l'EM sont

$$(E) : \ \check{\mathbf{x}}_{k+1} = \mathrm{E}\left[\mathbf{X}|\mathbf{y}, \bar{\mathbf{y}}_k\right] \tag{2.16}$$

$$(M) : \ \bar{\boldsymbol{\lambda}}_{k+1} = \text{argmax}_\lambda\{\lambda^T \mathbf{A}\check{\mathbf{x}}_{k+1} - \mathcal{F}^*(\lambda^T \mathbf{A})\}. \tag{2.17}$$

Ces deux étapes correspondent bien à la démarche que nous avons adoptée, équations (2.12) et (2.15). Par cette identification avec l'algorithme EM, on peut prévoir la convergence de la procédure. Notons cependant que l'étape délicate va être ici l'étape E, avec le calcul de la moyenne a posteriori pour lequel il faut, hors cas spécifique, échantillonner suivant la loi a posteriori. L'étape M correspond quant à elle au problème de maximum d'entropie sur la moyenne, pour lequel la solution est obtenue par un simple algorithme de gradient, d'autant plus rapide que l'on peut utiliser la solution obtenue à l'itération précédente comme condition initiale.

L'originalité de la démarche est ainsi d'introduire un a priori a maximum d'entropie sur la moyenne, dans une approche bayésienne, « réconciliant » ainsi les deux approches. La procédure résultante s'interprète ensuite simplement comme un algorithme EM dont les proprétés sont bien documentées.

Sur ce sujet, j'ai encadré sur quelques mois un travail de post-doctorat d'une étudiante tchèque, Pavla Holejšovská, qui, en dépit d'un investissement modeste, a obtenu des résultats numériques encourageants qui ont été présentés lors de conférences [6, 13, 7]. Il fait partie de mes projets de reprendre ce travail.

2.3 Maximum d'entropie avec contraintes aléatoires

Quittons maintenant le champ du maximum d'entropie sur la moyenne pour revenir à la problématique standard du maximum d'entropie : trouver la distribution la plus générale, la moins compromettante, compatible avec les contraintes. Le problème qui m'intéressait dans ce cadre était d'examiner ce que l'on peut dire lorsque les contraintes, justement, ne sont pas bien définies. Cette réflexion étant alimentée,

bien entendu, par l'expérience selon laquelle les données sont parfois bruitées. Un autre guide, il faut le reconnaître, était la connaissance de différents articles de physique définissant des "superstatistiques" [14, 15, 16, 17]. Dans ces travaux, les auteurs considèrent la loi d'équilibre d'un système, typiquement la distribution de Boltzmann-Gibbs – une distribution exponentielle, et imaginent que le paramètre naturel β est susceptible de varier selon une certaine distribution de probabilité. Dans ces conditions, on est face à une loi jointe, qu'il convient d'intégrer par rapport à la loi sur β pour obtenir la nouvelle description du système. Il est connu que l'on peut retrouver ainsi la distribution de Tsallis en choisissant une loi Gamma sur β. Ce fait, qui a suscité une abondante littérature en statistiques non-extensive, avait été noté depuis bien longtemps [18].

Dans le petit article [8], annexé au document en page 70, je considère des fluctuations non sur le paramètre de la loi de Boltzmann, mais sur les observables. Celles-ci servent classiquement de contraintes dans le cadre d'une approche à maximum d'entropie. Si $\bar{E}$ est la valeur de la contrainte (énergie moyenne), alors l'entropie correspondante s'écrit $S(\bar{E}) = \max_\beta\{-\beta\bar{E} - \log Z(\beta)\}$, et le paramètre β et la valeur moyenne sont deux variables duales reliées par une transformée de Legendre. Ces deux variables ne sont donc pas indépendantes. Aussi, lorsque les observables sont affectées par des fluctuations suivant une distribution $f_{\bar{E}}(\bar{E})$, le paramètre dual β est lui même affecté par des fluctuations suivant une loi $f_\beta(\beta)$ image de la précédente par la transformation. Ceci montre que l'on peut donc interpréter les fluctuations sur le paramètre naturel comme résultant d'une incertitude sur les contraintes.

À titre d'exemple, on considère le modèle canonique, avec une contrainte sur la valeur de l'énergie moyenne. Dans ce cadre, il est intéressant de considérer une loi Gamma $\Gamma(a, b)$

$$f_{\bar{E}}(\bar{E}) = \frac{b^a}{\Gamma(a)}\bar{E}^{a-1}\exp\left(-b\bar{E}\right)$$

comme modèle de fluctuations de cette énergie moyenne, soit une loi inverse-Gamma sur le paramètre β :

$$f_\beta(\beta) = \frac{b^a}{\Gamma(a)}\beta^{-a-1}e^{-\frac{b}{\beta}}.$$

Après intégration, il en résulte alors la distribution suivante

$$P(E) = \frac{2}{E}\frac{1}{\Gamma(a)\,\Gamma(3/2)}(bE)^{\frac{3/2+a}{2}}\,K_{-\frac{3}{2}+a}\left(2\sqrt{bE}\right). \tag{2.18}$$

où K_ν est une fonction de Bessel modifiée de seconde espèce et d'ordre ν. On souligne que cette distribution, dite K-distribution a trouvé des applications importantes dans le domaine de l'imagerie radar. On discute ensuite de sa caractérisation, de ses principales propriétés et de son comportement asymptotique (dont son caractère "heavy-tailed"). Les lecteurs pourront retrouver plus de détails dans l'article annexé.

2.4 Estimation de l'entropie

Pour pouvoir utiliser l'entropie directement dans des applications de traitement du signal ou de traitement de données, il peut être utile de pouvoir évaluer l'entropie attachée aux données considérées. Des estimateurs d'entropie sont présentés dans la synthèse [19] : on y trouve des estimateurs de type "plug-in" à partir d'estimées de la densité (typiquement à noyaux), des estimateurs reposant sur les statistiques d'ordre (dont l'estimateur de Vasiceck), et un estimateur fondé sur les plus proches voisins (estimateur de Leonenko, 1987). Dans un travail mené avec Christophe Vignat, nous avons envisagé une approche "signal" et nous nous sommes attachés à construire un estimateur de l'entropie de Shannon, qui possède des caractériques récursives, ce qui permet de l'employer « en ligne ». En outre, nous avons envisagé quelques applications et décrit l'apport d'un tel estimateur pour ces applications. Nous avons proposé deux estimateurs de l'entropie. Le premier repose sur une modélisation paramétrique des densités de probabilité et a fait l'objet d'une communication [20] et d'un article de revue [9]. Cet article est disponible en annexe page 76. Le second estimateur est un estimateur classique, reposant sur une technique de

type histogramme, pour lequel nous avons trouvé une formulation récursive sur le temps, et a été exposé dans une communication [21].

L'originalité du premier estimateur est de proposer d'assimiler la densité de probabilité à un spectre de puissance (normalisé), et d'utiliser un modèle autorégressif régularisé, cher à Jean-François Giovannelli, pour estimer cette densité de probabilité. Les exemples donnés dans l'article, que l'on retrouvera ici page 79 indiquent la pertinence et l'applicabilité du modèle. En réalité l'idée d'utiliser des méthodes d'estimation spectrale dans le cadre d'estimation de densités de probabilité avait été exposée un peu auparavant par Pagès & Lagunas d'une part et S. Kay d'autre part (mais les résultats numériques sont moins intéressants qu'avec des AR longs). Au passage, nous avons introduit un processus sous-jacent, défini par $W(n) = \exp(jX(n) + \phi)$ dont la densité spectrale de puissance $S_W(x)$ est précisément la densité de probabilité $f_X(x)$ de X (il suffit d'observer que sa fonction d'autocorrélation est la fonction caractéristique). À partir d'un modèle de la densité de probabilité sous la forme d'une fraction rationnelle, on est alors capable d'établir une expression explicite de l'entropie, en fonction des pôles de la fonction de transfert et des résidus associés ; ce qui n'est tout de même pas très pratique. En poursuivant avec une approche signal, on utilise simplement la relation de Plancherel-Parseval qui indique ici que

$$\hat{H}(X) = -\int_{-\frac{1}{2}}^{+\frac{1}{2}} S_W(x) \log_2 S_W(x) dx = -\sum_{k=-\infty}^{+\infty} R_W(k) C_W^*(k) \tag{2.19}$$

où R_W et C_W sont respectivement la fonction d'autocorrélation et cepstre de $W(n)$. Grace à la structure AR du signal, les séquences de corrélation et le cepstre peuvent être exprimés récursivement à partir des coefficients AR ou de la réponse impulsionnelle. Ainsi, il suffit d'estimer p coefficients de corrélation pour en déduire les coefficients AR, donc la réponse impulsionnelle, puis étendre les séquences de corrélation et de cepstre autant que souhaité. La formule (2.19) permet ensuite d'éviter le calcul explicite de la densité et l'intégration, en résumant le calcul de l'estimateur de l'entropie à un simple produit scalaire. De plus, on peut calculer récursivement la matrice de corrélation, en introduisant un facteur d'oubli, ce qui permet d'utiliser l'estimateur dans un cadre non stationnaire. On peut également alléger la charge calculatoire en utilisant un algorithme du gradient pour déterminer les paramètres AR plutôt que d'inverser une matrice de grande taille. Comparé aux estimées obtenues avec un estimateur par noyaux de la densité de probabilité, à l'estimateur de Vasciceck modifié, on observe que notre estimateur se comporte agréablement en termes de biais et de variance. La charge en calcul est également réduite. On a ensuite envisagé quelques applications pour lesquelles on peut utiliser directement un estimateur d'entropie : détection de ruptures dans un signal, déconvolution aveugle et séparation de sources. Dans les trois cas, l'utilisation de cet estimateur d'entropie se révèle pertinente. Signalons ici que nous retrouverons plus loin, en §4.3 page 39 le problème de détection ou de modification de la loi instantanée d'un signal, que nous avons également étudié en surveillant conjointement l'entropie et l'information de Fisher. On notera aussi que nous avons abordé le problème de déconvolution aveugle en nous appuyant sur le fait, conséquence de l'inégalité sur la puissance entropique (Entropy Power Inequality - EPI), que l'entropie du signal en sortie d'un système linéaire, alimenté par une séquence i.i.d. non gaussienne est toujours supérieure à l'entropie de l'entrée, si la norme du filtre est unitaire. Ce résultat se comprend intuitivement : sous une contrainte de puissance constante, le mélange des composantes d'un signal iid le rapproche de la gaussienne. Pour déconvoluer, on peut alors penser à ajuster les paramètres d'un filtre inverse de sorte à minimiser l'entropie de sortie. Le minimum absolu sera atteint lorsque l'on aura retrouvé la séquence i.i.d de départ.

Dans la même lignée, nous avons étudié un second estimateur, reposant sur un simple histogramme des données. Dans ce cadre, et dans la communication associée [21], nous avons proposé une interprétation des histogrammes comme l'estimation de la densité spectrale de puissance du « processus sous-jacent » à l'aide d'un banc de filtres ; donné une formulation récursive sur le temps d'un estimateur de l'entropie de Shannon fondé sur une analyse par histogramme, et enfin décrit une application pour la détection de changement de loi, et une application en séparation de sources.

2.5 Inégalité de convolution pour l'entropie de Rényi

Une difficulté posée par l'entropie de Shannon est que des résultats analytiques sont difficiles à obtenir dans le cas où l'on analyse des signaux subissant des transformations linéaires, comme par exemple dans la procédure de déconvolution évoquée dans le paragraphe précédent. Avec Christophe Vignat, nous avons cherché à exploiter la famille des entropies de Renyi, pour des applications de déconvolution.

Dans la communication [10], nous avons donné une inégalité sur l'entropie de Rényi qui étend (faiblement) l'inégalité de convolution pour l'entropie de Shannon (l'EPI). Cette inégalité repose sur une version étendue de l'inégalité de Young et sur l'aide de Franck Barthe. Nous avons montré que si $\mathbf{X}$ et $\mathbf{Y}$ sont deux vecteurs aléatoires indépendants, alors

$$H_r(\mathbf{X}+\mathbf{Y}) \geq \max(H_r(\mathbf{X}), H_r(\mathbf{Y})), \tag{2.20}$$

avec égalité si et seulement si $\mathbf{X}$ ou $\mathbf{Y}$ est un vecteur déterministe. À partir de cette inégalité, en considérant $\mathbf{Z}$ la sortie d'un système alimenté par une séquence de vecteurs iid $\mathbf{X}_i$, alors

$$H_r(\mathbf{Z}) \geq H_r(\mathbf{X}) + \frac{1}{2}\log|f_i|^2, \quad \forall i \tag{2.21}$$

avec égalité si et seulement si le filtre est un retard pur. Cette inégalité est à rapprocher de celle que l'on obtient pour l'entropie de Shannon :

$$H_1(Z) \geq H_1(X) + \frac{1}{2}\log\sum_i |f_i|^2, \tag{2.22}$$

avec égalité si et seulement si $X(n)$ est gaussien ou le filtre est un retard pur. En comparant les deux relations, on voit que l'inégalité pour l'entropie de Rényi reste plus faible, car elle fait apparaître la norme infinie du filtre et un seul de ses coefficients. Cependant le résultat n'est pas inintéressant puisqu'il permet d'obtenir des résultats explicites. Il est amusant de remarquer que le même résultat a été obtenu pratiquement simultanément (et en parfaite indépendance) par Hild, Erdogmus et Principe [22]. Il a été repris comme un résultat original dans plusieurs publications parmi lesquelles [23, 24, 25, 26, 27]. La dernière référence [27] met en garde sur un mauvais usage de l'entropie de Rényi en séparation de sources mais valide notre procédure de déconvolution.

En utilisant l'inégalité de convolution pour l'entropie de Rényi, on peut utiliser la procédure suivante : rechercher les coefficients d'un égaliseur tel que l'entropie en sortie du système global soit minimale. Il faut ajouter à cela une contrainte fixant un des coefficients, disons f_k à un (ou à l'identité) : $|f_k| = 1$. Cette contrainte évite que la minimisation ne conduise au filtre et à une sortie nulle pour laquelle on a effectivement égalité dans (2.21). Dans la mesure ou les f_k dépendent à la fois de l'égaliseur et du filtre inconnu, la contrainte précédente ne peut pas être assurée. Par contre, on peut fixer l'un des coefficients à une constante. Une solution est de fixer le premier coefficient de l'égaliseur, h_0, ce qui fixe le premier coefficient $f_0 = h_0 g_0$ de la réponse impulsionnelle globale. Dans le cas d'une entropie quadratique, et en utilisant un estimateur par noyaux, nous avons pu exprimer analytiquement l'entropie de Rényi, son gradient (et son hessien) par rapport aux coefficients matriciels d'un égaliseur (et les calculs correspondants furent formateurs). Des résultats de simulation nous ont montré que ce critère permettait, en dépit certainement de minima locaux, d'obtenir une déconvolution parfaite en aveugle, [10]. Terminons ce chapitre par une petite illustration : on considère un système MIMO avec $p = 2$ entrées et $q = 3$ sorties, avec la fonction de transfert suivante :

$$g(z) = \begin{bmatrix} 1 & -0.7 \\ 0.82 & 1 \\ 0.7 & 0.3 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & -0.3 \\ 0.1 & -0.2 \end{bmatrix} z^{-1} + \begin{bmatrix} 0.3 & -0.5 \\ 0.6 & -0.5 \\ 1 & 0.8 \end{bmatrix} z^{-2}.$$

Avec un égaliseur de longueur $L = 12$, la procédure précédente converge vers une solution intéressante. La réponse impulsionnelle globale du système, qui est très proche de l'identité, est donnée figure 2.1.

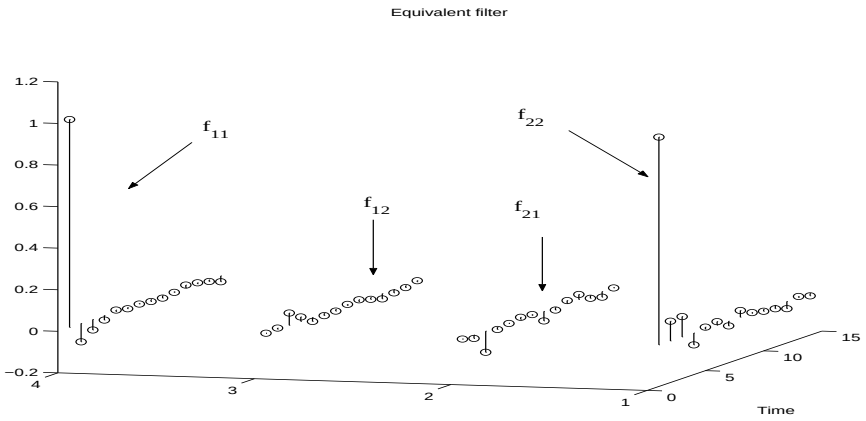


FIG. 2.1 – Réponses impulsionnelles équivalentes.

Chapitre 3

Entropies de Rényi-Tsallis

Le dernier paragraphe du chapitre précédent, sur une inégalité de convolution pour l'entropie de Rényi, nous permet une transition douce. Nous nous étions intéressés à l'entropie de Rényi pour examiner si, du point de vue pratique, on pouvait obtenir des résultats plus aisément manipulables qu'avec l'entropie de Shannon. Les entropies alternatives, telle que celle de Rényi, ont également des fondements axiomatiques qui en justifient l'utilisation. Parente de l'entropie de Rényi, l'entropie de Tsallis a suscité ces 20 dernières années un nombre important de recherches en physique statistique. L'une des raisons, au moins d'un point de vue pragmatique, de ce succès est que sa maximisation permet de faire apparaître des lois puissance, et que des lois de ce type semblent très importantes en pratique, voir par exemple la synthèse [28]. La thermodynamique associée, qui recouvre la thermodynamique habituelle comme cas particulier semble également en mesure de mieux expliquer certains phénomènes. En m'investissant sur ces questions, j'ai cherché à comprendre les raisons qui peuvent justifier ces entropies alternatives, quelles sont leurs propriétés remarquables, et comment on peut les utiliser. Il est connu par exemple que les distributions de Tsallis peuvent résulter d'une fluctuation selon une loi gamma du paramètre naturel d'une loi exponentielle. Nous avons discuté ce point en section 2.3 et proposé de relier ces fluctuations à des fluctuations de la contrainte.

En observant que les distributions de Tsallis apparaissent dans des contextes avec un équilibre déplacé ou perturbé, j'ai proposé un modèle qui permet d'interpréter les distributions de Tsallis comme résultant d'une procédure standard de maximum d'entropie (de Shannon), où l'on a ajouté une contrainte pour modéliser un déplacement de l'équilibre. Cette procédure, présentée ici en 3.2, a fait l'objet des publications [29, 30].

J'ai ensuite réalisé que les distributions de Tsallis sont présentes dans nombre de domaines, et coïncident avec la distribution de Pareto généralisée. Or cette distribution est la distribution limite qui décrit les queues de distribution en théorie des excès. Il s'agit probablement là d'une remarque importante qui peut expliquer une certaine ubiquité des distributions de Tsallis. Ce résultat a été présenté, avec Christophe Vignat, dans la référence [31]. Au passage, nous avons montré comment on pouvait connecter la distribution des excès, asymptotiquement, avec la solution d'un problème de maximum d'entropie [32]. Ces deux points sont présentés dans la section 3.3.

Une procédure de maximum d'entropie (de Rényi-Tsallis), ou de minimum de la divergence associée, telle qu'elle apparaît notamment comme résultante des modèles précédents, définit implicitement une entropie contractée, fonction de la contrainte et non plus de la distribution. Il est alors légitime de s'intéresser aux propriétés et caractérisations de ces "fonctionnelles entropiques". Ceci a fait l'objet du travail présenté dans la section 3.4 et dans la publication [33] placée ici en annexe.

3.1 Contexte

Dans son article de 1961 [34] Rényi présente un ensemble de postulats qui permet d'obtenir l'entropie de Shannon comme unique mesure d'information. L'un des postulats décrit que l'entropie de l'union

de deux distributions incomplètes est la moyenne arithmétique des entropies pondérée par les poids des distributions. En remplaçant dans ce postulat la moyenne arithmétique par une moyenne géométrique, Rényi obtient alors l'entropie de Rényi attachée à une distribution $\mathcal{P}$:

$$H_\alpha(\mathcal{P}) = \frac{1}{1-\alpha} \log \left(\sum_{k=1}^{n} p_k^\alpha \right). \tag{3.1}$$

Avec la même démarche, il donne également l'expression de l'information d'ordre α, $D_\alpha(\mathcal{P}||\mathcal{Q})$, dont nous reparlerons abondamment plus loin, qui est l'information qui apparaît lorsqu'une distribution $\mathcal{P}$ est remplacée par une distribution $\mathcal{Q}$.

Un autre résultat, apparemment bien peu connu et qu'il faudrait enseigner, qui justifie pleinement l'entropie de Rényi est un théorème de codage de source dû à Campbell (1965) [35]. Dans cette contribution, Campbell propose d'introduire une mesure de longueur des mots codes pénalisant les codes de longueur élevée. En effet, dans la définition classique, le coût associé à un mot code croît linéairement avec sa longueur. En notant D la taille de l'alphabet, n_i les longueurs élémentaires associées aux mots codes de probabilités p_i, Campbell définit alors

$$L(t) = \frac{1}{t} \log_D \sum_i p_i D^{t n_i} \tag{3.2}$$

qui pénalise la longueur de chacun des mots codes de manière exponentielle (il s'agit d'une transformation monotone simple de la longueur plus intuitive $\sum_i p_i D^{t n_i}$). Dans ces conditions, L.L. Campbell montre que

$$L(t) \geq H_\alpha, \tag{3.3}$$

où H_α est l'entropie de Rényi définie en (3.1), avec $\alpha = 1/(1+t)$. L'égalité est atteinte lorsque les longueurs des mots codes sont

$$n_i = -\log_D \left(\frac{p_i^\alpha}{\sum_i p_i^\alpha} \right), \tag{3.4}$$

où le terme de droite fait apparaître une "escort distribution" que nous retrouverons plus bas. Ce théorème étend ainsi de manière naturelle le théorème standard du codage de source, à nouveau par la substitution d'une moyenne arithmétique par une moyenne géométrique, et légitime l'entropie de Rényi. Resterait bien entendu à développer des codeurs à longueur (au sens de Campbell) variable. [1] [2]

L'entropie de Tsallis [37] a été introduite en 1988 dans le champ de la physique statistique, à l'origine pour la description des multifractales. Elle est définie par

$$S_\alpha(f_X) = \frac{1}{1-\alpha} \left(\int f_X(x)^\alpha \, dx - 1 \right) \tag{3.5}$$

où α est un réel positif et f_X une distribution univariée. Comme dans le cas de l'entropie de Rényi, l'entropie de Shannon est obtenue, par la règle de L'Hospital, dans le cas limite $\alpha = 1$.

Notons au passage, car cela semble largement ignoré, que l'entropie de Tsallis est suggérée dans l'article initial de Rényi [34, eq. 4.20, p 561] :

> It is clear from the above proof that instead of the quantities (4.4) we could have used the analogous sums
>
> (4.20) $\quad \sum p_{jk}^{(n)} f\left(\frac{p_{jk}^{(n)}}{p_k} \right)$
>
> where $f(x)$ is any function such that $xf(x)$ is strictly convex. Thus for instance we could have taken $f(x) = x^{\alpha-1}$ with $\alpha > 1$ or $f(x) = -x^{\alpha-1}$ with $0 < \alpha < 1$.

[1] Campbell a introduit une pénalisation exponentielle de la longueur des mots codes. Une autre manière d'introduire une fonction de coût, ou longueur, pénalisant les mots-codes de grande taille associée à des probabilités faibles serait de déformer la distribution de probabilité pour tenir davantage compte des indices de probabilités (initialement) faible. Dans cet esprit il serait intéressant de rechercher un résultat de type codage de source pour une moyenne généralisée définie comme $L(\alpha) = \sum_i n_i P_i$, avec P_i une "escort distribution" $P_i = p_i^\alpha / \sum_k p_k^\alpha$.

[2] Depuis la rédaction du mémoire, un travail dans le sens indiqué dans la note précédente a été initié, et dont les résultats ont été présentés dans [36].

Les distributions de Tsallis peuvent être obtenues par maximisation de l'entropie de Tsallis (3.5), sous contrainte de moyenne et de normalisation. Bien entendu, comme l'entropie de Rényi et de Tsallis sont liées par une transformation monotone, la maximisation de l'une fournit le même argument que la maximisation de l'autre. Deux types de contraintes ont été considérées dans la littérature : la moyenne statistique standard $\mathrm{E}[X] = m$, et une moyenne généralisée $\mathrm{E}_\alpha[X] = m$, qui est prise par rapport à une 'escort' ou 'zooming distribution' $p(x)^\alpha / \int p(x)^\alpha dx$. Ces deux contraintes conduisent à un résultat de la forme

$$f_X(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma}x\right)^{-\frac{1}{\gamma}-1} \text{ pour } x_F > x \geq 0, \tag{3.6}$$

où σ et γ sont respectivement les paramètres d'échelle et de forme. Pour $\gamma < 0$, la densité a un support fini, avec $x_F = -\sigma/\gamma$ et un support infini dans le cas contraire. Dans le premier cas, l'exposant est $1/(\alpha - 1)$ et $\sigma = 1/\beta$; dans le second cas, l'exposant est $1/(1-\alpha)$ et $\sigma = 1/\beta(2-\alpha)$. Pour $\gamma = 0$ ($\alpha = 1$), la solution se réduit à la distribution exponentielle standard $f_X(x) = 1/\sigma \exp(-x/\sigma)$. Comme je l'ai déjà souligné, un des intérêts des distributions de Tsallis est de pouvoir rendre compte d'un comportement en loi puissance, qui est souvent rencontré dans des applications.

Depuis l'introduction de l'entropie de Tsallis en physique statistique, un grand nombre d'études, et un nombre non moins considérable de publications (plus de 1000 références dans la bibliographie tenue par Tsallis), se sont penchées sur l'étude des entropies alternatives et des thermodynamiques associées. C. Tsallis a publié très récemment un livre de synthèse [38]. En effet, il apparaît que dans le cas de certains systèmes complexes, la thermodynamique standard semble insuffisante pour une description complète. Les applications incluent les problèmes de turbulence, les statistiques des rayons cosmiques, l'économétrie, etc. Dans nombre d'exemples, les résultats d'expérimentation et les résultats numériques coïncident très bien avec une description par une distribution de Tsallis.

3.2 L'entropie de Rényi-Tsallis comme solution d'un problème standard de maximum d'entropie

Le principe du maximum d'entropie peut être vu, interprété, comme une conséquence du théorème de Sanov, comme nous l'avons déjà indiqué dans le chapitre 2, section §2.1, eq. 2.3. L'essence de celui-ci est que dans l'ensemble de toutes les distributions issues d'une distribution q, compatibles avec une contrainte, celle qui devient absolument prépondérante est la distribution la plus proche de q au sens de l'information de Kullback-Leibler. Dans ces conditions, et dans la mesure où les distributions de Tsallis sont différentes des distributions obtenues par une approche de maximum entropie classique, celles-ci seront avec les mots de Grendar [39] *will be asymptotically improbable*. Compte-tenu du succès rencontré dans les applications, il est légitime de se demander si on peut trouver des justifications probabilistes pour l'entropie de Tsallis ou pour sa maximisation qui conduit aux distributions de Tsallis. Un essai de synthèse a été présenté lors du workshop Iwap 2008 [40]. Une première approche est proposée ici. L'idée sous jacente a été présentée au colloque MaxEnt2006 [29], puis après une métamorphose dans la publication [30], qui est annexée ici page 88.

Le point de départ pour rechercher une justification de l'apparition des distributions de Tsallis est de noter qu'elles semblent apparaître dans le cas d'un équilibre de Boltzmann-Gibbs modifié, perturbé, ou déplacé. On peut essayer de rendre compte de cela en introduisant dans la formulation du maximum d'entropie une contrainte supplémentaire dont l'effet est de déplacer l'équilibre. Cet équilibre déplacé peut être imaginé comme résultant d'un compromis entre deux distributions, R et Q. Au lieu de sélectionner la distribution la plus proche de la référence Q sous une contrainte de moyenne, on recherche une distribution P^* proche, simultanément, des deux références. Une telle distribution sera localisée quelque part entre R and Q. Par exemple, on considèrera un système global composé de deux sous-systèmes caractérisés par deux distributions distinctes. La distribution d'équilibre sera atteinte pour un intermédiaire entre les deux distributions de référence, et l'observable pourra être, selon les cas, soit la moyenne sur un sous système, soit la moyenne sous la distribution d'équilibre. Un tel modèle peut rendre compte d'un processus de fragmentation : un système $\Sigma(A, B)$ se fragmente en A, avec une distribution R et

B avec une distribution Q, et le système complet est vu avec la distribution d'équilibre P^*. On peut aussi voir une transition de phase : un système quitte l'état Q vers un état R, et présente une distribution intermédiaire P^*.

Ceci peut être posé de la manière suivante : trouver une distribution P^* telle que la divergence de Kullback-Leibler à Q, $D(P||Q)$ soit minimale, sous la contrainte d'équilibre entre R et Q, par exemple $D(P||Q) - D(P||R) \leq \theta$, ou de manière équivalente, une contrainte $D(P||R) \leq \theta'$. On peut remarquer que la première contrainte correspond à un log-rapport de vraisemblance moyen. Ceci s'écrit

$$\left\{ \begin{array}{c} \min_P \ D(P||Q) = \min_P \ \int P(x) \log \frac{P(x)}{Q(x)} dx \\ s.c. \ \theta = D(P||Q) - D(P||R) = \int P(x) \log \frac{R(x)}{Q(x)} dx \leq \theta. \end{array} \right. \tag{3.7}$$

En fait, ce petit problème a été considéré, et résolu, par Kullback [41, page 39] comme une illustration de son théorème général sur la minimisation de $D(P||Q)$ sous contrainte, et l'on obtient :

$$P^*(x) = \frac{R(x)^\alpha Q(x)^{1-\alpha}}{\int R(x)^\alpha Q(x)^{1-\alpha} dx}, \tag{3.8}$$

expression qui n'est rien d'autre que l'"escort distribution" de la physique nonextensive (avec en plus ici l'apparition du terme en Q). Le paramètre α est simplement le paramètre de Lagrange associé à la contrainte. La distribution P^* est simplement la moyenne géométrique (pondérée) entre R et Q et réalise un compromis, ajusté par α, entre les deux références. Par dualité de Lagrange, on a ici l'égalité entre le problème initial et son dual :

$$\left\{ \begin{array}{c} \min_P \ D(P||Q) \\ s.c. \ \theta = D(P||Q) - D(P||P_1) \end{array} \right. = \sup_\alpha \left(\alpha\theta - \log \left(\int P_1(x)^\alpha Q(x)^{1-\alpha} dx \right) \right). \tag{3.9}$$

On peut alors noter que dans cette dernière relation, le terme $\log \left(\int P_1(x)^\alpha Q(x)^{1-\alpha} dx \right)$, qui est strictement concave en α, est directement proportionnel à la divergence de Rényi. En d'autres termes, la solution du problème initial, $D(P^*||Q)$ s'exprime simplement à partir de la divergence de Rényi d'ordre α :

$$D(P^*||Q) = \alpha\theta - (\alpha - 1) D_\alpha(R||Q). \tag{3.10}$$

Il n'est pas difficile de vérifier que la solution satisfait une relation de Pythagore :

$$D(P||Q) = D(P||P^*) + D(P^*||Q) \tag{3.11}$$

pour toute distribution P telle que $D(P||Q) - D(P||R) = \int P \log \frac{R}{Q} = \theta$. La solution optimale P^* s'interprète alors comme la projection orthogonale sur la ligne joignant R et Q, ce qui minimise à la fois la distance à R et à Q.

On peut utiliser différentes variantes de l'argumentation, en considérant par exemple un problème de maximisation de l'entropie, avec une contrainte sur la distance à une loi de référence Q. C'est l'approche adoptée dans [30], où on imagine que Q représente la distribution des excès. On cherche alors une distribution à maximum d'entropie dont la « queue » est à une distance (de Kullback) θ de la distribution de départ.

Remarques

Il est fort intéressant également de noter que la contrainte sur le log rapport de vraisemblance n'est pas totalement déraisonnable. Dans son mémoire de doctorat, Marcel-Paul Schützenberger [42] introduit en pages 57 à 61, une information qu'il appelle information de Wald [43] qui inclut notamment le log- rapport de vraisemblance moyen entre deux hypothèses. On la rencontre en analyse séquentielle. Ainsi, notre problème initial s'interprète comme la recherche de la distribution la plus proche d'une distribution de référence, au sens de la divergence de Kullback, sous contrainte que l'information de Wald-Schützenberger

soit bornée (ou fixée). Notons également que dans ce mémoire tout à fait précurseur, Marcel-Paul Schützenberger propose une démarche unificatrice des « théories des informations ». Entre autres, il définit, page 65, la fonction génératrice des cumulants c'est-à-dire la log-fonction de partition, rencontrée dans le chapitre 2 dans le problème de maximum d'entropie standard, comme « pseudo-information ».

Poursuivant sur l'interprétation du log rapport de vraisemblance, on peut encore noter que les décodeurs itératifs modernes produisent des sorties souples qui sont des log-rapports de vraisemblance. On comprend donc que les probabilités d'erreurs doivent alors pouvoir être obtenues par l'application d'un principe de grande déviation sur la somme des log-rapports de vraisemblance, i.e. leur moyenne empirique. Comme le principe de grandes déviations fait intervenir comme fonction de taux le minimum de l'information de Kullback sous la contrainte, ici de log-rapport de vraisemblance, on voit émerger l'entropie de Rényi de paramètre α, comme en (3.10). Ce paramètre α est le paramètre de Lagrange associé à la contrainte et lié à celle-ci par une opération de dérivation. Bien qu'assez fier de cette interprétation, je dois renvoyer les lecteurs vers l'article [44] à l'auteur duquel revient la primeur de cette remarque. Curieusement, il ne me semble pas l'avoir vu reprise par ailleurs.

Observables

Les valeurs observables sont habituellement la valeur moyenne sous une certaine distribution. Ici, selon le point de vue, l'observable pourrait être soit la moyenne sous R, la distribution d'un sous système, ou encore sous la distribution d'équilibre P^*. Ainsi, on peut chercher à minimiser plus avant la divergence de Kullback-Leibler $D(P||Q)$, une fois la contrainte de log-vraisemblance moyenne atteinte, en ajustant R de sorte à vérifier la contrainte de moyenne. Ceci peut s'écrire

$$K = \left\{ \begin{array}{l} \min_{P_1} \left\{ \begin{array}{c} \min_P \; D(P||Q) = \min_P \; \int P(x) \log \frac{P(x)}{Q(x)} dx \\ \text{s.c.} \quad \theta = \int P(x) \log \frac{R(x)}{Q(x)} dx \\ \text{s.c.} \;\; m = E_R[X] \text{ or } m = E_{P^*}[X] \end{array} \right. \end{array} \right. , \tag{3.12}$$

En prenant en compte (3.9), on obtient alors

$$K = \sup_{\alpha} \left[\alpha\theta - \left\{ \begin{array}{c} \max_R (\alpha - 1) D_\alpha(R||Q) \\ \text{s.c.} \quad m = E_{P_1}[X] \text{ or } m = E_{P^*}[X] \end{array} \right. \right] \tag{3.13}$$

qui revient simplement à la recherche d'un extremum de la divergence de Rényi sous contrainte de moyenne. On notera également que la moyenne sous la distribution P^* n'est rien d'autre que la moyenne généralisée utilisée en « statistiques nonextensives ». Bien entendu, les solutions correspondantes sont les distributions de Tsallis que nous avons indiquées en (3.6).

3.3 Distributions de Tsallis et distribution des excès

C'est en étudiant les domaines d'apparition, ou d'utilisation, de modèles de distribution de lois puissance, et notamment à la lecture de [45] que je vis apparaître une possibilité de connexion des « statistiques non extensives » et des résultats de la théorie des excès (Extreme value theory). En effet, on s'aperçoit que la distribution de Tsallis a exactement la même forme (3.6) que la distribution de Pareto généralisée (GPD) qui apparaît en théorie des excès. Le point intéressant est que ces distributions de Pareto généralisée sont utilisées hors du champ de la physique statistique pour modéliser des distributions à queues lourdes. Les exemples sont nombreux, allant de la théorie de la fiabilité, au trafic dans les réseaux, à l'hydrologie, la climatologie, la géophysique, la science des matériaux, l'imagerie radar et l'économie.

La pertinence, sinon toujours une justification, de ce modèle de distribution de Pareto généralisée dans ces applications peut être liée au théorème de Balkema-de Haan-Pickands [46, 47] qui indique que la distribution des excès au dessus d'un seuil élevé est souvent bien approchée par une GPD. Le résultat de Pickands et Balkema-De Haan relie l'appartenance au domaine d'attraction d'une des distributions extrême au comportement de la queue de la distribution considérée. Il montre de plus que la convergence

vers une des trois distributions limite n'a lieu que si la queue de la distribution, vue au delà d'un seuil de plus en plus élevé, est décrite par une GPD, avec une erreur aussi faible que l'on veut. Plus spécifiquement, on considère une variable aléatoire X, un seuil u et on note X_u la variable conditionnelle représentant les excès au dessus du seuil u

$$X_u = X - u | X > u.$$

La fonction de survie correspondante s'écrit alors

$$\bar{F}_{X_u}(x) = \Pr(X > x + u | x > u) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)}, \ x \geq 0. \tag{3.14}$$

Le théorème de Pickands indique que la loi de X appartient au domaine de convergence d'une distribution extrême si et seulement si on peut trouver γ et σ tel que

$$\lim_{u \to x_\infty} \inf_{\sigma > 0} \| \bar{F}_{X_u}(x) - \bar{G}_{\gamma,\sigma}(x) \|_\infty = 0$$

où $\bar{G}_{\gamma,\sigma}$ est la fonction de survie associée à la densité de la GPD en (3.6) :

$$\bar{G}_{\gamma,\sigma} = \left(1 + \frac{\gamma}{\sigma} x\right)_+^{-\frac{1}{\gamma}} \tag{3.15}$$

avec la notation $(x)_+ = \max(0, x)$. Le cas $\gamma > 0$ correspond au domaine de Fréchet, avec des lois à queues lourdes (Cauchy, t, Pareto), le cas $\gamma < 0$ à des lois à support borné (x appartient à l'intervalle $[0, -\sigma/\gamma]$) du domaine de Weibull, et enfin le cas $\gamma = 0$ correspond à une distribution exponentielle (domaine de Gumbel – exponentielle, gamma, normale, K, etc).

Jusqu'alors, les distributions de Tsallis, qui sont des GPD, étaient justifiées soit par le recours au principe de Maximum d'entropie de Tsallis, soit par fluctuations du paramètre de la loi de Boltzmann (suivant une loi gamma - de manière analogue à ce que nous avons décrit en section 2.3). En notant que ces distributions sont des GPD, et en nous appuyant sur le résultat de Pickands, nous voyons donc que ces distributions peuvent également être interprétées statistiquement comme résultant d'un seuillage ou d'une série de seuillages. C'est le point de vue original que nous avons exposé et illustré dans l'article [31] qui est annexé ici en pages 94 à 104.

Dans cet article, nous avons également fait un parallèle entre le théorème de Pickands et le théorème central limite et souligné la propriété de stabilité des GPD par seuillage (au sens ou une GPD reste une GPD avec le même exposant). Nous avons présenté un certain nombre de procédures d'estimation des paramètres de la GPD. En effet, pour l'estimation des paramètres de distributions à queues lourdes, il convient d'être précautionneux, ce qui n'est pas forcément toujours le cas dans les applications en statistiques nonextensives [48, 49]. Enfin, nous avons repris brièvement les résultats que nous avons exposés à la conférence ISIT [32] dans lesquels nous établissons une autre connexion avec le maximum d'entropie de Tsallis.

Dans cet autre travail, on commence par montrer que la GPD peut être obtenue comme solution d'un problème de maximum d'entropie de Rényi-Tsallis, sous des contraintes de moyenne et de normalisation. Dans un second temps, on considère les distributions appartenant au domaine d'attraction de Fréchet, et on caractérise les normes 1 et q (la norme q étant proportionnelle à l'entropie de Tsallis) de la fonction de survie associée à la variable d'excès $X - u | X > u$. En utilisant une normalisation appropriée, on définit une nouvelle variable, qui peut être aussi simple que X_u/u, dont la moyenne et les normes 1 et q de la fonction de survie convergent vers des valeurs constantes. Le même travail est fait pour une sous-classe des distributions du domaine de Gumbel. On montre alors que la distribution des excès est une distribution de Tsallis. Dans le cas Fréchet, on identifie le paramètre q comme une fonction simple du paramètre de la famille de Fréchet, afin que la moyenne, les moments d'ordre 1 et q coïncident avec ceux de la variable des excès normalisée. Dans la mesure où la solution d'un problème de maximum d'entropie sous contraintes est unique, on obtient que la variable des excès obéit asymptotiquement à une loi de Pareto généralisée, qui est solution d'un problème de maximum d'entropie de Rényi-Tsallis.

3.4 Fonctionnelles entropiques dérivées de la divergence de Rényi

Dans la référence [29], reprise dans le paragraphe 3.2, nous avons vu que l'on peut ramener un problème de minimisation de l'information de Kullback, sous contrainte de log-vraisemblance et de moyenne, à un problème de minimisation de la divergence de Rényi. Le minimum de l'information sous contrainte de moyenne m est bien sûr une fonction de cette moyenne, qui est une entropie dans l'espace des observations. Dans le cas de l'information de Kullback, ceci correspond au principe de contraction qui amène aux fonctions de taux (entropie de niveau 1) dans la théorie des grandes déviations. Dans la lignée des travaux précédents, il est tout à fait naturel de se pencher sur les propriétés de la minimisation de la divergence de Rényi et des fonctionnelles entropiques associées. L'article [33] qui est annexé ici en pages 108 à 125 reprend les résultats qui ont été obtenus.

On a considéré une minimisation de la divergence de Rényi sous une contrainte de moyenne standard, mais aussi, comme en physique statistique, sous une contrainte de moyenne généralisée. Celle-ci est, on l'a vu plus haut, la moyenne prise par rapport à l'"escort distribution"

$$P^*(x) = \frac{P(x)^\alpha Q(x)^{1-\alpha}}{\int_{\mathcal{D}} P(x)^\alpha Q(x)^{1-\alpha} dx}. \tag{3.16}$$

Celle-ci se réduit d'ailleurs à la distribution P quand α tend vers un, et la moyenne généralisée redevient la moyenne habituelle. Le problème de minimisation de la divergence de Rényi $D_\alpha(P||Q)$ d'une distribution P à une distribution Q, sujet à une contrainte de normalisation et à une contrainte de moyenne classique (C) ou généralisée (G) s'écrit comme suit :

$$\mathcal{F}_\alpha^{(C \text{ resp. } G)}(m) = \left\{ \begin{array}{rl} & \min_P D_\alpha(P||Q) \\ \text{s.c.} & (C) \quad m = E_P[X] \\ \text{ou} & (G) \quad m = E_{P^*}[X] \\ & \text{et} \quad \int_{\mathcal{D}} P(x)dx = 1 \end{array} \right. \tag{3.17}$$

où $\mathcal{F}_\alpha^{(C)}(m)$ et $\mathcal{F}_\alpha^{(G)}(m)$ sont les entropies de niveau 1 associées à la divergence de Rényi pour les contraintes classique et généralisée.

Dans le travail présenté dans l'article annexé [33], on étudie les formes et propriétés des solutions au problème ci-dessus, sous les deux types de contraintes. Les valeurs prises à l'optimum, vues comme fonctions de la contrainte, définissent des fonctionnelles entropiques $\mathcal{F}_\alpha^{(.)}(m)$ associées à chaque choix de mesure de référence Q et indexées par le paramètre α. Il me semble que cette étude est complètement originale.

On commence par établir la forme exacte des distributions P qui réalisent le minimum de l'information de Rényi dans le terme de droite de (3.17), et on montre que ces distributions, définies sur un sous-ensemble $\mathcal{D}$ de $\mathbb{R}$ (avec notamment $\mathcal{D}$ tel que l'intégrale soit finie et la densité de probabilité toujours réelle positive), sont de la forme

$$P_\nu(x) = \frac{[1 + \gamma\,(x - \bar{x})]^\nu}{Z_\nu(\gamma, \bar{x})} Q(x). \tag{3.18}$$

La normalisation $Z_\nu(\gamma, \overline{x})$ est une fonction de partition vérifiant

$$Z_\nu(\gamma, \overline{x}) = \int_{\mathcal{D}} [1 + \gamma(x - \overline{x})]^\nu \, Q(x) dx \tag{3.19}$$

Plus spécifiquement, on montre que dans le cas de la contrainte (C), on a $\nu = \xi = \frac{1}{\alpha - 1}$ et $\overline{x} = E_P[X] = E_\xi[X]$; et dans le cas de la contrainte généralisée, on a $\nu = -\xi = \frac{1}{1-\alpha}$ et $\overline{x} = E_{P_G^*}[X] = E_{-(\xi+1)}[X]$, la moyenne généralisée. Il est important de voir que $\overline{x}$ est dans les deux cas une moyenne statistique et non la contrainte m, et par suite une fonction du paramètre γ. C'est uniquement dans le cas du paramètre optimal γ^* que l'on a bien entendu $\overline{x} = m$.

On liste quelques propriétés des distribution P_ν en (3.18) et de leurs fonctions de partition (3.19). En particulier on montre que la divergence de Rényi, et par suite les fonctionnelles entropiques $\mathcal{F}_\alpha^{(.)}(m)$, s'expriment simplement comme le logarithme de ces fonctions de partition. On établit ensuite des propriétés générales de ces fonctionnelles, dont le fait qu'elles sont toujours non-négatives et présentent un minimum unique en m_Q, la moyenne sous la référence Q. De plus, on montre que $\mathcal{F}_\alpha^{(C)}(m)$ est strictement convexe, pour $\alpha \leq 1$. On montre également que les fonctionnelles $\mathcal{F}_\alpha^{(.)}(x)$ vérifient une symétrie spécifique : $\mathcal{F}_\alpha^{(C)}(x) = \mathcal{F}_{\frac{1}{\alpha}}^{(G)}(x)$. À partir de ces fonctionnelles, on définit aussi une divergence dans l'espace des moyennes possibles, cette divergence se réduisant à une divergence de Bregman lorsque $\alpha \to 1$.
Dans la mesure où les distributions optimales (3.18) sont définies implicitement, puisqu'elles dépendent de leur moyenne, la détermination des paramètres n'est pas évidente (il existe des techniques itératives, mais qui supposent que la solution est un attracteur). Pour résoudre ceci, on définit deux fonctionnelles dont la maximisation fournit le paramètre γ, pour les deux types de contraintes. On vérifie que les valeurs γ^* du paramètre γ des distributions optimales solution de (3.17) sont les minimum des maximiseurs de

$$D_C(\gamma) = -\log Z_{\xi+1}(\gamma, m) \tag{3.20}$$
$$D_G(\gamma) = -\log Z_{-\xi}(\gamma, m) \tag{3.21}$$

où les deux fonctions de partition sont convexes, mais éventuellement définies sur plusieurs intervalles. Les fonctionnelles entropiques associées $\mathcal{F}_\alpha^{(.)}$ sont alors données par

$$\mathcal{F}_\alpha^{(C \text{ resp. } G)}(m) = D_{C \text{ resp. } G}(\gamma^*). \tag{3.22}$$

Remarque - C'est dans le cas discret et pour $\nu < 0$ (ce qui correspond à la contrainte standard (C) avec $\alpha \in (0, 1)$) , que la fonction de partition $Z_\nu(\gamma, m)$ présente des singularités sur son support, et n'est définie que sur des segments. Dans ce cas de figure, $-\log Z_\nu(\gamma, m)$ peut présenter plusieurs maxima. Cela indique que pour une moyenne donnée, il existe plusieurs distributions de la famille P_ν qui admettent cette moyenne. Réciproquement, on peut aussi montrer que pour un paramètre gamma donné (et γ est homogène à un « inverse de température »), il peut exister plusieurs moyennes (états) associées. On peut même définir une fonction potentiel dont les maxima fournissent les valeurs de ces moyennes. Ces différentes solutions semblent correspondre à des distributions dont le support est différent, ou s'étend. Ce phénomène qui reflète le fait que la transformation $\gamma \leftrightarrow \bar{x}(\gamma)$ n'est pas bijective, peut être imaginé comme une transition de phase. Il y a sur ces points un travail à effectuer.

On présente ensuite 4 cas particuliers de mesure de référence $Q(x)$: des distributions uniforme et exponentielle dans le cas de systèmes à états continu, puis une distribution de Bernoulli et une distribution de Poisson. Dans chaque cas, on calcule la fonction de partition et les fonctions duales associées, et on cherche à déterminer une expression des fonctionnelles $\mathcal{F}_\alpha^{(.)}(x)$. On vérifie que l'on obtient des fonctionnelles connues (Burg, Fermi-Dirac, Kullback), lorsque α tend vers 1. Dans tous les cas, on fournit des évaluations numériques des fonctionnelles entropiques. Les figures Fig. 3.1 et 3.2, extraites de l'article, illustrent les comportements que l'on peut obtenir. L'exploitation de ces fonctions pour des problèmes de sélection, de résolution de problèmes inverses peut faire l'objet d'une poursuite de ce travail.

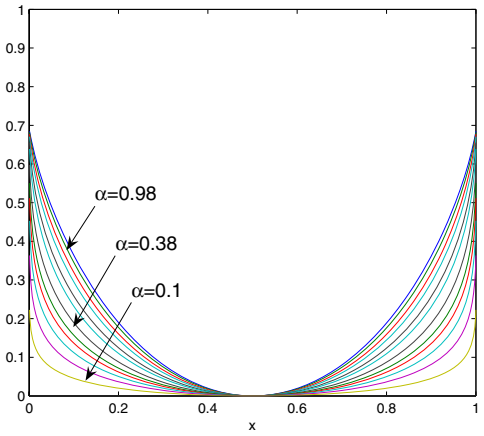


FIG. 3.1 – Fonctionnelle entropique $\mathcal{F}_\alpha^{(C)}(x)$ dans le cas d'une mesure de Bernoulli et $\alpha \in (0,1)$.

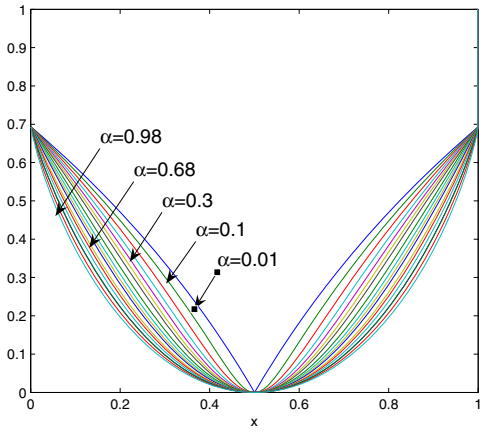


FIG. 3.2 – Fonctionnelle entropique $\mathcal{F}_\alpha^{(G)}(x)$ dans le cas d'une mesure de Bernoulli et $\alpha \in (0,1)$.

Chapitre 4

Information de Fisher

L'information de Fisher est d'une importance fondamentale en théorie de l'estimation, notamment par le théorème de Cramér-Rao sur la variance d'un estimateur. Elle apparaît également dans des problèmes d'inférence et d'interprétation en physique, par le principe d'*Extreme Physical Information* défendu par Roy Frieden, avec de nombreux exemples [50, 51, 52]. L'information de Fisher est encore utilisée comme un outil de caractérisation de signaux et systèmes [53, 54], avec des applications en géophysique, [55, 56, 57], en biologie [58], en reconstruction [59, 60] en traitement du signal [61, 62].

Avec Christophe Vignat, nous nous sommes intéressés à trois problèmes liés à l'information de Fisher, et ce chapitre reprend les contributions correspondantes. La première concerne la transformation de l'information de Fisher par des systèmes singuliers [63]. La seconde étudie l'interaction de l'information de Fisher et de l'entropie de Shannon et introduit le fameux « plan d'information de Fisher-Shannon » [64]. Enfin, dans un travail récent, nous étudions les distributions d'information de Fisher minimale, avec une variance fixée et sous contrainte de support borné [65].

4.1 Quelques éléments de contexte

L'information de Fisher est une mesure de l'information portée par la loi d'une variable aléatoire X sur un paramètre, généralement noté θ. Si $f(x,\theta)$ est la vraisemblance, la fonction score est définie par

$$\phi_\theta(X) = \frac{\partial}{\partial\theta}\log f(X;\theta) = \frac{1}{f(\theta;X)}\frac{\partial f(\theta;X)}{\partial\theta}.$$

et l'information de Fisher associée à θ est la variance de la fonction score :

$$I(\theta) = \mathrm{E}\left\{\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2\middle|\,\theta\right\},$$

où l'on a utilisé le fait que la moyenne de la fonction score est nulle.

Si $T(X)$ est une statistique pour θ, alors

$$I_T(\theta) \le I_X(\theta),$$

avec égalité uniquement si $T(X)$ est suffisante.

L'inégalité de Cramér-Rao fournit une borne sur l'erreur quadratique moyenne. Si $T(X)$ est un estimateur de θ, alors on a toujours

$$\mathrm{E}\left[(T(X)-\theta)^2\right] \ge \frac{(1+B(\theta)')^2}{I(\theta)} + B(\theta)^2, \tag{4.1}$$

avec $B(\theta) = \mathrm{E}\left[T(X)-\theta\right]$. Dans le cas non biaisé, l'inégalité se réduit à

$$\mathrm{Var}\left\{T(X)\right\} \ge \frac{1}{I(\theta)}. \tag{4.2}$$

On sait également que l'information de Fisher apparaît dans l'expression de l'information de Kullback-Leibler entre deux densités de probabilité, où le paramètre de l'une est le perturbé de l'autre. Dans le cas d'une perturbation infinitésimale, on obtient que l'information de Kullback-Leibler devient une forme quadratique pondérée par l'information de Fisher. L'information de Fisher est alors une métrique sur la variété des distributions de probabilité paramétrées.

Dans le cas multivarié, si $\boldsymbol{\theta}$ est le vecteur $\boldsymbol{\theta} = \left[\theta_1, \theta_2, \ldots, \theta_N\right]$, alors l'information de Fisher est une matrice de terme générique

$$\left(I\left(\boldsymbol{\theta}\right)\right)_{i,j} = \mathrm{E}\left[\frac{\partial}{\partial\theta_i}\ln f(X;\theta)\frac{\partial}{\partial\theta_j}\ln f(X;\theta)\right]. \tag{4.3}$$

Toujours dans le cas non-biaisé, l'inégalité de Cramér-Rao s'écrit alors $\mathrm{Var}\left\{\mathbf{T}(\mathbf{X})\right\} \geq I(\boldsymbol{\theta})^{-1}$, où l'inégalité s'entend comme le caractère défini non-négatif de la différence des deux matrices.

L'information de Fisher est définie pour une famille paramétrique, contrairement à l'entropie, qui est définie pour toute distribution. On peut en fait définir une famille paramétrée associée à toute distribution en considérant le paramètre de localisation, et définir l'information de Fisher associée à la famille $f(x-\theta)$. Dans ce cas, la dérivation par rapport à θ est équivalente à la dérivation par rapport à x, et on obtient alors

$$I(X) = \mathrm{E}\left\{\left[\frac{\partial}{\partial x}\ln f(X)\right]^2\right\} = \int_{-\infty}^{+\infty} f(x)\left[\frac{\partial}{\partial x}\ln f(x)\right]^2 \mathrm{d}x. \tag{4.4}$$

Il s'agit de *l'information de Fisher de la distribution*. Lorsque x est vectoriel et le paramètre de translation scalaire, l'information de Fisher attachée à la distribution est scalaire. Si le paramètre de translation est vectoriel, alors c'est une matrice d'information de Fisher de la distribution qui apparaît en (4.4). Dans le cas d'un paramètre d'échelle σ, on obtient $I(X) = \int_{-\infty}^{+\infty}\left(1 + xp(x)'/p(x)\right)^2 p(x)\mathrm{d}x$, en $\sigma = 1$.

C'est essentiellement à l'information de Fisher sur la distribution (4.4) que nous nous sommes intéressés (jusqu'ici). Pour cette information, on dispose d'une inégalité de convolution, la Fisher Information Inequality, FII [66] :

$$I(X+Y)^{-1} \geq I(X)^{-1} + I(Y)^{-1}, \tag{4.5}$$

qui est aussi valable dans le cas où X et Y sont des vecteurs. En utilisant cette inégalité et le théorème de De Bruijn, on peut obtenir l'inégalité de la puissance entropique, l'Entropy Power Inequality EPI. Si on pose

$$N(\mathbf{X}) = \frac{1}{2\pi e}\exp\frac{2}{d}H(\mathbf{X}) \tag{4.6}$$

où $H(\mathbf{X})$ est l'entropie de Shannon et d la dimension de X, alors l'EPI s'écrit

$$N(\mathbf{X}+\mathbf{Y}) \geq N(\mathbf{X}) + N(\mathbf{Y}). \tag{4.7}$$

On peut encore compléter ces inégalités par des relations d'incertitude (isoperimetric inequality), cf [66]

$$\frac{1}{d}I(\mathbf{X})N(\mathbf{X}) \geq 1 \text{ et } N(\mathbf{X})|\mathbf{I}(\mathbf{X})|^{\frac{1}{d}} \geq 1 \tag{4.8}$$

où $I(\mathbf{X})$ est l'information de Fisher de $\mathbf{X}$ par rapport à un paramètre de translation scalaire et $\mathbf{I}(\mathbf{X})$ est la matrice d'information de Fisher. On notera que si on applique un facteur d'échelle sur le vecteur aléatoire, $\mathbf{Y} = \mathbf{A}\mathbf{X}$, l'inégalité de droite reste vraie.

Enfin, pour un paramètre de translation, l'inégalité de Cramér-Rao est

$$I(\mathbf{X})\mathbf{K}(\mathbf{X}) \geq \mathbf{1}, \tag{4.9}$$

où $\mathbf{K}$ est la matrice de covariance de la loi de X.

4.2 Inégalités sur l'information de Fisher

Les premiers résultats sur la transformation de l'information de Fisher ont été donnés dans les années 60 par Stam [67] et Blachman [68]. Il s'agit en particulier de l'inégalité

$$I^{-1}(\sum_i a_i X_i) \geq \sum_i a_i^2 I^{-1}(X_i).$$

En 1993, Papathanasiou [69] donna une importante série d'inégalités sur l'information de Fisher, avec des applications à la caractérisation de la normalité. En parallèle et indépendamment, Zamir [70] établit ce même résultat et discuta les cas d'égalité [71, 72]. Il s'agit de l'inégalité

$$I^{-1}(\mathbf{Y}) \geq \mathbf{A} I^{-1}(\mathbf{X}) \mathbf{A}^T \tag{4.10}$$

où $\mathbf{A}$ est une matrice de rang lignes plein, $\mathbf{X}$ est un vecteur aléatoire à composantes indépendantes, et $\mathbf{Y} = \mathbf{AX}$. Une autre inégalité est trouvée dans ce travail :

$$I(\mathbf{X}) \geq \mathbf{A}^T I(\mathbf{Y}) \mathbf{A}. \tag{4.11}$$

Ces deux inégalités constituent la version de FII multivariée, donnée par Zamir. On note par ailleurs que dans le cas inversible, on a simplement $J(\mathbf{AX}) = \mathbf{A}^{-T} J(\mathbf{X}) \mathbf{A}^{-1}$, ce qui permet de retrouver le membre de droite dans (4.10). Cette borne fournit une fonction de contraste possible pour les problèmes de déconvolution et séparation de sources dans le cas convolutif.

Cependant, les démonstrations fournies dans les articles de Zamir et complétées dans son rapport technique [72] sont assez ardues (représentation conditionnelle de la fonction score et extension de l'inégalité de Stam au cas vectoriel), et plus encore pour ce qui concerne l'étude des cas d'égalité. Nous avons pu trouver une preuve très simple de la FII de Zamir, en nous appuyant simplement sur une inégalité sur les matrices partitionnées. Donnons l'idée de la démarche. On considère la factorisation LDL d'une matrice symétrique, définie non-négative, partitionnée selon $\mathbf{U} = [\mathbf{A} \;\; \mathbf{B}; \mathbf{C} \;\; \mathbf{D}]$, avec $\mathbf{B} = \mathbf{C}^T$ et $\mathbf{D}$ carrée de rang plein. La condition de définie non-négativité entraîne $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} \geq 0$, avec égalité si $\mathbf{U}$ a le même rang que $\mathbf{D}$. Il ne reste plus qu'à identifier les matrices qui permettent de retrouver ((4.10) et ((4.11) : dans le premier cas on utilise la matrice de covariance de $[\phi_X(\mathbf{X})\phi_Y(\mathbf{Y})]^T$, et dans le second cas celle de $[\phi_Y(\mathbf{Y})\phi_X(\mathbf{X})]^T$, où les ϕ sont les fonctions score. Enfin pour conclure, il faut utiliser le fait que $\mathrm{E}\left[\phi_X(\mathbf{X})\phi_Y(\mathbf{Y})^T\right] = \mathbf{A}^T I(\mathbf{Y})$, une égalité qui demande un peu de réflexion. Les cas d'égalité se déduisent ensuite à partir de la condition de rang, et on a relié ceux-ci aux concepts de pseudoinverse et de gaussianité.

Par ailleurs, dans sa démonstration originale, Zamir utilise l'expression de la fonction score de l'observation $\mathbf{Y} = \mathbf{AX}$ en fonction de la moyenne conditionnelle de $\mathbf{X}$. On cherche à exprimer $\phi_X(X)$ en fonction de $\phi_Y(\mathbf{Y})$, sachant que $\mathbf{Y} = \mathbf{AX}$, avec $\mathbf{A}$ non inversible. Nous avons formulé le problème comme un problème d'estimation au sens MMSE : $\mathrm{argmin}_w \mathrm{E}\left[|\phi_X(\mathbf{X}) - w(\mathbf{Y})|^2\right]$ sous la contrainte $\mathbf{Y} = \mathbf{AX}$. On obtient alors que l'estimateur s'écrit $\mathrm{E}\left[\phi_X(X)|Y\right] = A^T \phi_Y(Y)$, retrouvant ainsi les résultats de Zamir, avec même l'économie d'une hypothèse. Ces différents résultats ont été présentés à ISIT, [73], où nous avons rencontré Ram Zamir. Un article [63], a ensuite été publié dans Journal of Inequalities in Pure and Applied Mathematics. Cet article est annexé en pages 130 à 137. Notons enfin que ces résultats ont été repris dans une entrée de Mathworld http://mathworld.wolfram.com/FisherInformationMatrix.html.

4.3 Plan d'information de Fisher-Shannon

Dans le travail précédent, nous nous étions intéressés à l'extension de l'inégalité sur l'information de Fisher (4.5). Nous nous sommes ensuite attardés sur la relation d'incertitude (4.8), dans le cas scalaire cette fois :

$$I(X)N(X) \geq 1. \tag{4.12}$$

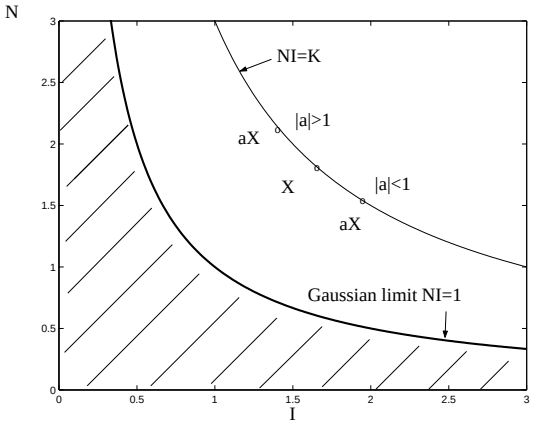


FIG. 4.1 – Plan d'information de Fisher-Shannon

En remarquant que par une transformation d'échelle a, on a $I(aX) = a^{-2}I(X)$ et $N(aX) = a^2N(X)$, on voit que le produit $I(X)N(X)$ est invariant par cette transformation. Ces deux propriétés, relation d'incertitude et invariance montrent que l'information de Fisher et l'entropy power sont liées. Nous avons alors introduit la notion de plan d'information de Fisher-Shannon et proposé de caractériser un phénomène physique (ou plutôt la loi sous-jacente) par une position dans le plan d'information de Fisher-Shannon. Dans ce plan, la transformation d'échelle précédente se traduit par l'évolution sur une courbe $I(X)N(X) = K$, où K est une constante. Ceci est illustré sur la figure Fig. 4.1.

Nous avons étudié deux familles de densité de probabilité : les Student-t, qui sont les distributions de Tsallis obtenues sous contrainte de variance, et les gaussiennes généralisées qui sont également souvent utilisées en physique (et même en signal pour modéliser les statistiques des coefficients d'ondelettes). Dans les deux cas, nous avons donné les expressions de l'entropie et de l'information de Fisher, et nous avons montré qu'il existe une bijection entre les paramètres de la distribution et la position dans le plan Fisher-Shannon.

Une évolution des paramètres des lois se traduit alors par une trajectoire dans le plan Fisher-Shannon. Il est alors possible de construire des trajectoires arbitraires dans le plan Fisher-Shannon ; nous avons ainsi construit une trajectoire en escalier, où l'information de Fisher reste constante tandis que l'entropy power croît, avant de basculer sur une croissance de l'information de Fisher tandis que l'entropy power est fixée. Ceci signifie donc que l'étude d'un seul des deux indicateurs peut être insuffisante et qu'aucun des indicateurs n'est plus adéquat que l'autre. Une autre trajectoire arbitraire est une trajectoire circulaire dans le plan d'information. Cependant, la représentation reste incomplète : on peut par exemple trouver des distributions de variances différentes qui partagent la même coordonnée dans le plan Fisher-Shannon, et il peut être alors judicieux de s'intéresser à une représentation Fisher-Shannon-variance (notons que dans le cas des deux familles que nous avons étudié, un point de même coordonnée en Fisher-Shannon-variance correspond à une loi gaussienne qui est l'intersection de nos deux familles). Bien entendu, on pourrait aussi trouver des distributions différentes qui partageraient une même coordonnée dans la représentation Fisher-Shannon-variance : la bonne caractérisation, c'est la connaissance de toute la loi. La proposition est ici d'une caractérisation partielle, mais qui présente une certaine cohérence en terme d'information. Un article issu de ce travail a été publié dans Physics Letters A, et est ici proposé en annexe, pages 142 à 148, où l'on pourra retrouver des représentations de différentes réalisations suivant les trajectoires décrites ci-dessus.

Cette idée de représentation conjointe a rencontré un certain succès, notamment comme outil d'analyse de systèmes d'atomes [74, 75, 76], de systèmes électroniques [77], (dans ces deux cas, on possède des modèles analytiques des lois de comportement et on peut ainsi comparer des systèmes à l'aide de cet indicateur simple), pour l'analyse de processus d'ionisation [56, 78]. D'autres plans d'information ont été également proposés : Cramér-Rao-Variance [79], Fisher-Rényi [80], ou encore Fisher-Complexité [78, 81].

4.4 Minimum de l'information de Fisher

Une des caractéristiques de l'information de Fisher est bien sûr de faire intervenir explicitement, dans sa définition même, la dérivée de la densité de probabilité. Tout problème d'extrémisation de l'information de Fisher va alors faire apparaître une équation différentielle du second ordre. C'est pour cette raison, au moins du point de vue opérationnel, que l'on peut retrouver la plupart des équations de la physique à partir de l'information de Fisher. Ceci est discuté notamment dans les livres de Frieden [52, 82, 83] et dans de nombreux articles, où celui-ci définit un principe d'extrémisation de l'information EPI Extreme Physical Information, qui préconise d'extrémiser le « flot d'information », la différence de deux informations de Fisher associées respectivement à la source et aux données.

Dans les travaux exposés ci-dessus, nous avons étudié quelques aspects relatifs à l'inégalité sur l'information de Fisher FII (4.5), en examinant l'évolution de l'information lors de la transformation d'un vecteur aléatoire par un système singulier. Nous nous sommes ensuite intéressé à la relation d'incertitude (4.8) et avons introduit la notion de plan d'information de Fisher-Shannon pour l'analyse de signaux. Une autre inégalité fondamentale sur l'information de Fisher est bien entendu l'inégalité de Cramér-Rao (4.9).

Il est bien connu que l'inégalité de Cramér-Rao est saturée par la distribution gaussienne. Ainsi, la distribution de variance fixée qui minimise l'information de Fisher est la gaussienne. Ceci est vrai sur $\mathbb{R}$. Cependant, dans nombre de cas pratiques on sait que la variable d'intérêt appartient à un sous ensemble de $\mathbb{R}$ (positivité, normalisation, acquisition avec un support fini par exemple). Il est donc intéressant de rechercher, comme modèle de la distribution des données ou pour améliorer l'inégalité de Cramér-Rao, les distributions compatibles avec les données – nous choisissons ici une contrainte de variance, présentant une information de Fisher minimale. Ceci revient à rechercher et caractériser les distributions de support limité, dont l'information de Fisher (relative au paramètre de localisation), soit la plus faible, ce qui indique que le paramètre correspondant, par l'inégalité de Cramér-Rao, est le plus délicat à estimer. Cette étude a été également motivée par la lecture de l'article de Elke Uhrmann-Klingen [84] paru dans *Sankhya : The Indian Journal of Statistics* (1995), qui a analysé les propriétés générale des solutions sur un support compact, mais sans obtenir d'expression explicite. Elke Uhrmann-Klingen a eu la gentillesse de nous envoyer sa thèse de doctorat [85] (1992 - en allemand).

Dans le travail, qui m'a bien occupé ces derniers mois, nous avons obtenu les expressions explicites des densités probabilité d'information de Fisher minimale avec une variance fixée, sur $\mathbb{R}^+$ et sur un intervalle, on a étudié leur comportement et celui de l'information de Fisher associée. Ces résultats sont présentés dans un article accepté très récemment dans la revue *Information Sciences* [86] et annexés ici entre les pages 152 et 162. On considère ainsi le problème variationnel

$$I(\sigma^2) = \inf_f \left\{ I[f] : \mathrm{Supp}[f] = \mathcal{S}, f \in \mathcal{D} \text{ and } \mathrm{Var}[f] = \sigma^2 \right\} \tag{4.13}$$

où $\mathcal{S}$ désigne le support de la distribution. Une des difficultés rencontrées est que bien que $I[f]$, l'information de Fisher attachée à la distribution, soit convexe, la contrainte $\mathrm{Var}[f] = \sigma^2$ ne définit pas un ensemble convexe. En posant $f = u^2$ et en utilisant le calcul de variations, on obtient l'équation d'Euler-Lagrange

$$u''(z) - \left(\alpha + \beta z^2\right) u(z) = 0 \tag{4.14}$$

qui est une équation différentielle parabolique, établie par Weber pour l'étude des solutions de l'équation de Laplace pour des coordonnées cylindriques.

Les solutions de l'équation de Weber peuvent être exprimées à l'aide des fonctions parabolique cylindriques $D_\nu(x)$ et $D_{-\nu-1}(ix)$ [87]. On peut également se ramener à l'équation différentielle de Whittaker par changement de variable, dont les solutions s'expriment à partir des fonctions de Whittaker $M_{\lambda,\mu}(z)$ et $M_{\lambda,-\mu}(z)$, ou encore à l'aide de la fonction $W_{\lambda,\mu}(z)$ de Whittaker qui est reliée aux précédentes. Certaines de ces fonctions peuvent s'exprimer comme une série, ce qui permet d'étudier assez simplement les cas limite.

En prenant pour support les réels positifs, on a alors obtenu comme solution la famille

$$u_{\alpha,\beta}(z) \propto D_{-\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}((2\sqrt{\beta})^{\frac{1}{2}} z) \tag{4.15}$$

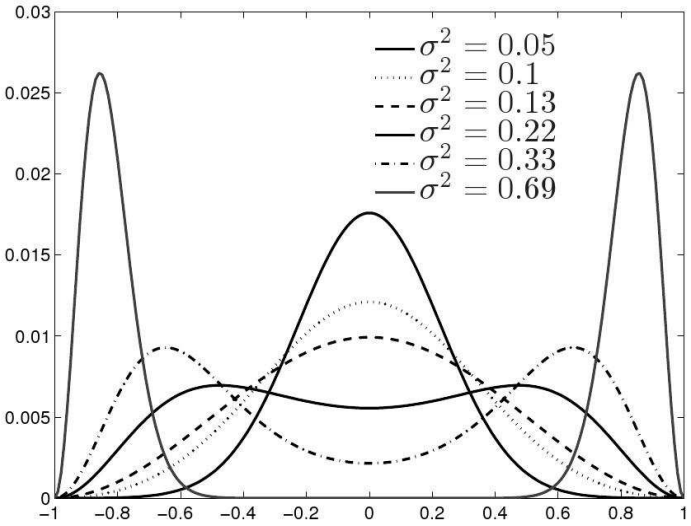


FIG. 4.2 – Densités de probabilité $f(x) = u(x)^2$, d'information de Fisher minimale sur l'intervalle $[-1, 1]$

où le dénominateur sert à normaliser à un la densité $f(z) = u(z)^2$. Pour ces densités, on a établi que

$$I_\nu(\sigma^2) = \frac{K_\nu}{\sigma^2}, \tag{4.16}$$

avec $K_\nu = I_{\nu,1} V_{\nu,1}$, où on a fixé $\xi = 1$. On a ensuite recherché la valeur ν^* de ν qui minimise K_ν de sorte à sélectionner le représentant de la famille d'information de Fisher minimale. Nous avons pu déterminer analytiquement les expressions de l'information de Fisher et de la variance, ce qui a permis d'étudier K_ν en fonction de ν. On a ainsi obtenu que $\nu^* = 1$ et ainsi que la distribution qui minimise l'information de Fisher sur $\mathbb{R}^+$, avec une contrainte de variance, est

$$f_\xi(x) = \sqrt{\frac{2}{\pi}} \xi^{\frac{3}{2}} x^2 \exp\left(-\frac{\xi x^2}{2}\right), \tag{4.17}$$

la chi-distribution à trois degrés de liberté, avec $\xi = \left(3 - \frac{8}{\pi}\right)/\sigma^2$. On obtient alors que le produit Fisher-variance est $K_1 = 9 - 24/\pi \approx 1.3606$. Au passage, on en déduit donc que sur $\mathbb{R}^+$, on a toujours

$$I[f] \geq \frac{K_1}{\sigma^2}. \tag{4.18}$$

Pour un support compact (ici $[-1, 1]$), nous avons obtenu l'expression analytique de la solution :

$$u(x) = \frac{\frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}}, -\frac{1}{4}}(\sqrt{\beta} x^2)}{\left(\int_{-1}^{1} \frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}}, -\frac{1}{4}}(\sqrt{\beta} x^2) \mathrm{d}x\right)^{\frac{1}{2}}}. \tag{4.19}$$

où $\sqrt{\beta}$ est le premier zéro $M_{-\frac{\alpha}{4\sqrt{\beta}}, -\frac{1}{4}}$. Cette dernière condition imposant l'annulation de la densité sur les frontières du domaine. Malheureusement, il n'existe pas de solution analytique donnant ces zéros, et les paramètres doivent être obtenus numériquement. La figure 4.2 donne quelques exemples du comportement de cette solution pour différentes contraintes de variance. On note que deux classes de solutions apparaissent : des solutions unimodales et d'autres pour une variance plus élevée, qui sont bimodales.
En fait, on a montré que la fonction $I(\sigma^2)$ est une fonction strictement convexe de σ^2, avec un minimum (forcément unique) en un certain σ_*, et que pour $\sigma < \sigma_*$, alors $\beta > 0$; pour $\sigma > \sigma_*$, $\beta < 0$ et finalement $\sigma = \sigma_*$, $\beta = 0$; où β est le paramètre de Lagrange de l'équation différentielle (4.14). Dans le cas $\beta < 0$, l'argument de la solution (4.19) devient complexe et cela provoque l'apparition des deux modes. Les cas

limite, pour $\sigma \to 0$ et $\sigma \to 1$ sont respectivement une densité gaussienne et deux masses sur les frontières de l'intervalle. Finalement, on obtient aussi facilement l'expression de la densité d'information de Fisher minimale sur l'intervalle, correspondant à la variance σ_* :

$$f(x) = \cos\left(\pi x/2\right)^2 \tag{4.20}$$

$$\text{avec } \sigma_*^2 = \frac{1}{3} - \frac{2}{\pi^2} \text{ et } I(\sigma_*^2) = \pi^2. \tag{4.21}$$

Des prolongations de ce travail, sont possibles, dans le cas multivarié par exemple, ou encore pour l'analyse des distributions conjointement difficiles, au sens de l'information de Fisher, pour le paramètre de localisation et le paramètre d'échelle.

On a vu que l'on peut soupçonner une « géométrie » autour des lois de Tsallis, qui étendent la notion de famille exponentielle. Cependant, lorsque l'on étudie le développement au second ordre des divergences de Rényi et Tsallis, on obtient l'information de Fisher standard, comme dans le cas de l'information de Kullback-Leibler : la géométrie différentielle associée semble inchangée. Il n'est toutefois pas illégitime d'étudier la correspondance distribution $\to$ escort distribution et de s'intéresser alors à l'information de Fisher associée aux escort distributions, définie par :

$$I_\alpha[f] = \frac{\alpha^2}{\int f(x)^\alpha \mathrm{d}x} \int f(x)^\alpha \left(\frac{df}{dx}\right)^2 \mathrm{d}x. \tag{4.22}$$

Il est facile de voir que l'on a une inégalité de Cramér-Rao reliant cette information de Fisher et la variance généralisée. En introduisant l'entropy power associé à l'entropie de Rényi selon

$$V_\alpha[f] = \left[\int f(x)^\alpha \mathrm{d}x\right]^{\frac{2}{d}\frac{\alpha'}{\alpha}} \tag{4.23}$$

avec $1/\alpha + 1/\alpha' = 1$, cf [88] et la définition classique dans le cas $\alpha = 1$, il est également possible de vérifier que le produit $V_\alpha[f].I_\alpha[f]$ est invariant par changement d'échelle. Existe-t-il une relation d'incertitude associée ? En tout état de cause, on a le lien suivant [89] entre l'information de Fisher généralisée et l'entropie de Rényi :

$$I_\alpha[f] \geq (\alpha - 1)^2 \left(\frac{\partial H_\alpha}{\partial \theta}\right)^2. \tag{4.24}$$

Une autre question serait de déterminer la forme des densités à minimum d'information de Fisher, sous contrainte de variance ou sous contrainte de variance généralisée. Enfin, on peut noter que les distributions de Tsallis, les Student-t et r, vérifient l'équation différentielle $f^{2\alpha-1}(x)f''(x) = [ax^2 + b]$ simple et étudier les problèmes qui feraient apparaître une équation différentielle de ce type.

Chapitre 5

Problèmes d'architecture Radio

La demande des usagers, la force commerciale qui conduit chacun à imaginer regarder des images animées en prenant le métro, les nouveaux usages (localisation GPS, lecture mp3, internet mobile, photographie, vidéo, etc) conduisent à un développement extraordinaire de la complexité, des applications et des systèmes qui doivent faire face à de multiples schémas de modulation, protocoles, bandes de fréquence multiples. Bien entendu les applications citées ci-dessus demandent une très forte intégration et des capacités de traitement du signal très importantes. Les usagers entendent aussi disposer de « multifonctions » (le terme est de Rufin dans son livre Globalia [90]) légers en poids, en prix et en consommation. En ce qui concerne l'aspect consommation, une des clés est de rechercher des architectures permettant d'utiliser un amplificateur dans un mode le plus efficace possible. Dans cet ordre d'idée, j'ai participé à un travail avec I. Jamin et C. Berland sur une évolution de l'architecture EER (Envelope Elimination and Restoration), qui est commenté en §5.1. Une autre clé va être d'intégrer au maximum la partie des composants analogiques radiofréquence (RF). Il semble que ce soit cependant une tâche particulièrement complexe, et que s'il est certainement possible de développer des composants analogiques hautement intégrés, le coût en étude, réalisation et consommation ne soit pas raisonnable. Une autre possibilité est de chercher à développer des alternatives numériques (avec un coût de production plus faible) aux composants analogiques. C'est le chemin qui a été initié notamment chez Texas Instrument par l'équipe de R. Staszewski avec le développement d'un "*digital RF processor*" [91], qui permet en outre d'aller vers des systèmes reconfigurables. Dans un travail avec C. Joubert, G. Baudoin, T. Divel, nous avons apporté quelques contributions à l'étude des boucles à verrouillage de phase entièrement numérique. Ces contributions sont abordées section 5.2. Avec C. Berland, nous pensons aussi qu'il faut profiter du traitement du signal, ainsi de plus en plus intégré dans l'émetteur (et le récepteur) pour compenser les défauts des composants analogiques, voire pour relâcher les contraintes sur ces composants. Ceci sera discuté plus avant dans le chapitre suivant.

La montée en fréquence nécessite aussi des composants alternatifs aux composants actuels. Ainsi, avec P. Guillot, C. Berland et P. Philippe, nous travaillons à la conception d'un oscillateur stabilisé à très haute fréquence mettant en jeu un oscillateur BAW (Bulk Acoustic Wave), et pouvant servir d'alternative aux systèmes à quartz ou évitant l'emploi d'une PLL. Ce travail est rapidement présenté section 5.3.

D'un autre côté, si on veut essayer de faire intervenir le traitement numérique du signal au plus tôt dans la chaîne de traitement, se pose le problème d'un échantillonnage précoce du signal. Avec A. Lesellier, O. Venard et O. Jamin, nous débutons une étude sur les solutions pour un échantillonnage et traitement de signaux très large bande dans le cadre d'une application de réception de télévision numérique.

Dans ces différents travaux, il s'agit de travaux conjoints dans lesquels les compétences de mes différents collaborateurs, seniors ou étudiants sont essentielles. Corinne Berland notamment m'a fait bénéficier de ses connaissances et entraîné sur ses thématiques d'architecture EER. C'est par l'interaction recherchée entre différentes visions que nous avons pu avoir quelques idées.

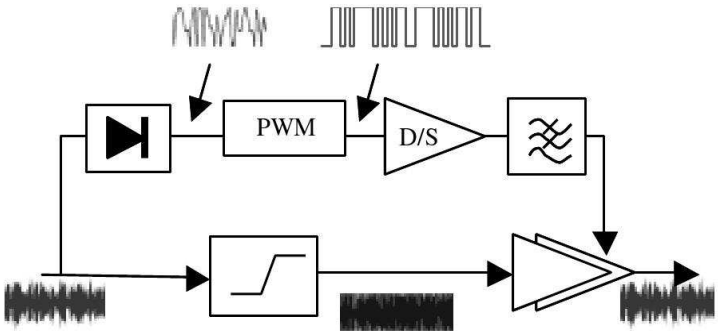


FIG. 5.1 – Architecture EER : version classique

5.1 Une architecture d'émetteur pour les modulations à enveloppe variable

Les systèmes modernes de radiocommunication, du fait des débits très importants qui sont visés, reposent sur des schémas de modulation où les modulations de type QAM sont préférées aux simples modulations de phase ou de fréquence. Alors que le standard 3GPP met en jeu des QPSK et 16QAM, des débits plus élevés, dans les applications de type WLAN ou DVB, sont atteints avec des modulations à multiporteuses. La contrepartie des débits élevés est que la dynamique de l'enveloppe du signal s'accroît, la bande passante s'élargit, ce qui pose des contraintes pour la conception et la réalisation de la partie analogique du dispositif. En effet, dans la conception d'un émetteur, il est impératif de rechercher à la fois l'efficacité et la linéarité. La variation de l'enveloppe, combinée aux conversions AM/AM et AM/PM de l'amplificateur de puissance dégradent l'EVM (Error Vector Magnitude), le spectre du signal transmis, et l'ACPR (Adjacent Channel Power Ratio). Des méthodes de linéarisation doivent être mises œuvre. Ces méthodes peuvent s'intéresser à corriger les défauts de linéarité des composants, notamment de l'amplificateur, où changer le problème en convertissant le signal en un signal équivalent à enveloppe constante. C'est l'idée initiale de l'EER (Envelope Elimination and Restoration) introduite initialement par Kahn [92]. Il s'agit d'une solution analogique où le signal est séparée en un signal d'enveloppe et un signal de phase, ce qu'on appelle également décomposition polaire. Le signal de phase est amplifié par un amplificateur de type classe C et l'enveloppe est restaurée en sortie du dispositif en modulant l'alimentation de l'amplificateur, voir la figure Fig. 5.1.

L'architecture classique de l'EER peut être modifiée de sorte à ce que la restauration intervienne par multiplication avant l'amplificateur de puissance. Plus exactement, on sépare l'enveloppe et la phase, et on code l'enveloppe par une modulation à largeur d'impulsions et on multiplie l'enveloppe ainsi codée par les signaux I et Q. On obtient alors un signal qui alimente un amplificateur fonctionnant en mode commuté. La variation d'enveloppe est restaurée après l'amplificateur par filtrage passe-bande, comme illustré Fig. 5.2.

Avec C. Berland, nous réfléchissions au problème de détermination d'une forme d'onde optimale pour coder l'entrée à l'aide d'une modulation à largeur d'impulsions PWM. J'ai suggéré d'utiliser un codage Sigma-Delta qui permet d'obtenir des performances comparables à l'aide de dispositifs existants. D'une part cette modification a pu permettre d'améliorer les performances par rapport au codage PWM, et d'autre part nous avons compris que cette modification, en utilisant un $\Sigma\Delta$ à sorties signées $\pm a$ permettait de reconstruire un simple signal à enveloppe constante portant à la fois les informations d'enveloppe et de phase du signal initial : nous avions isolé une procédure permettant de transformer n'importe quel signal en un signal à enveloppe constante, sans perte d'information. Plus encore, la restauration du signal initial se fait simplement par filtrage passe-bande. Du fait de la propriété d'enveloppe constante, le choix d'un amplificateur de puissance n'est plus limité aux amplificateurs commutés. Le schéma de principe est présenté figure Fig. 5.3.

Cette architecture a été étudiée et implantée lors du travail de thèse d'Isabelle Jamin, en convention Cifre

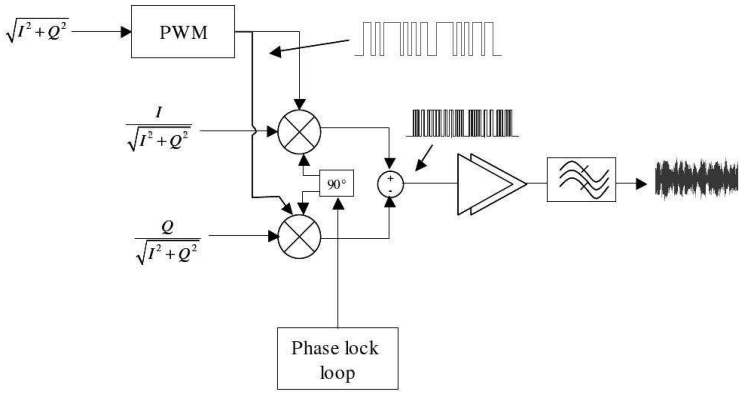


FIG. 5.2 – Nouvelle architecture EER, version commutée :

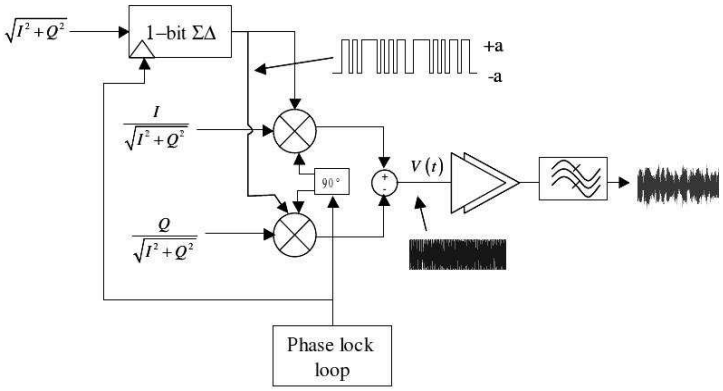


FIG. 5.3 – Nouvelle architecture EER à modulateur ΣΔ symétrique : version enveloppe constante

avec STMicroelectronics Crolles pour la la réalisation matérielle de cette architecture. L'architecture et ses possibilités d'extension a fait l'objet d'un brevet, en versions nationale et internationale [93] (2006). Dans l'article associé à ce travail, annexé page 166, nous présentons cette architecture, qui peut être adoptée pour tout type de modulation, sur une large gamme de fréquences et avec une grande efficacité et linéarité. Nous illustrons les performances de la solution à PWM, qui transforme le signal initial en un signal commuté, souffre de termes d'intermodulation et peut présenter des performances insuffisantes. Le second schéma, à modulateur $\Sigma\Delta$ permet de transformer le signal initial en un signal à enveloppe constante. Il met en jeu un modulateur de très haute fréquence. Le choix des filtres passe-bande est analysé, et on montre que la solution permet d'obtenir les performances escomptées en terme de spectre de sortie et d'EVM, pour une modulation sévère comme l'OFDM.

5.2 Boucle à Verrouillage de phase numérique

Pour le développement des systèmes modernes de communication, une des possibilités, comme indiqué dans l'introduction, est d'aller vers des architectures entièrement numériques. Dans cette direction, j'ai coencadré, avec G. Baudoin (directrice de thèse), la thèse Cifre de C. Joubert en convention Cifre avec St Microelectronics, (RF Expertise Center - Cellular Terminal Division), sur l'analyse et l'optimisation d'architectures fondées sur une boucle à verrouillage de phase numérique. L'intitulé exact du sujet était : « Contribution à l'analyse d'une boucle à verrouillage de phase à traitement de signal numérique : stabilité et cycles limites, sources de bruit de phase. ».

En effet, une architecture (et même plusieurs) de boucle à verrouillage de phase entièrement numérique a été proposée par R. Staszewski et son équipe (Texas Instrument). Une synthèse des approches et travaux de Staszewski et al. est présentée dans [94], ainsi que chapitre 10 du livre [91]. Cette architecture était très récente au début de ce travail, la publication « complète » datant de 2005 [95] (mais les éléments étaient disponibles avant dans différents articles de conférence). L'idée du travail de thèse est d'analyser cette architecture, de proposer éventuellement des aménagements, et de l'adapter pour un fonctionnement à plus haute fréquence, pour arriver jusqu'à la réalisation d'un prototype.

Nous avons en particulier proposé un modèle comportemental complet du système, qui permet de le simuler avec une charge en calcul très raisonnable. L'article de congrès [96] présente ce modèle. Nous avons également développé des algorithmes rapides permettant d'alléger l'implantation [97], et calculé explicitement la dsp du bruit de phase correspondant à la structure [98]. Un prototype d'un émetteur FM employant cette PLL et implantant sur silicium les algorithmes développés a été réalisé et validé. Nous décrivons ici les principaux résultats obtenus et annexons au document, page 171, une version longue et plus détaillé de l'article [97].

Un bloc diagramme de la PLL numérique est présenté Figure 5.4. Un oscillateur contrôlé numériquement, le DCO (digitally controlled oscillator) permet d'implanter cette PLL de manière complètement numérique. Les accumulateurs de phase sont utilisés pour compter les périodes de la référence et de la sortie du DCO. Une horloge FS synchronisée avec Fdco, sous-échantillonne la sortie du DCO de sorte que les comparaisons des deux phases soient réalisées avec la même horloge.

Une ADPLL de haute précision est obtenue en utilisant une erreur de phase fractionnaire. Cette erreur est proportionnelle au retard entre les signaux Fdco et Fref [notations étranges de chez TI, mais ce sont bien des signaux temporels]. Le TDC (Time to Digital Converter) est utilisé pour convertir ce retard en une quantité numérique, avec une résolution Δ égale au plus petit retard de propagation au travers d'une porte inverseuse, constituant élémentaire du TDC. Le mot de commande FCW (Frequency Command Word) est donné comme entrée de l'accumulateur de phase et permet de régler la fréquence de sortie selon

$$FDCO = FCW \times FREF.$$

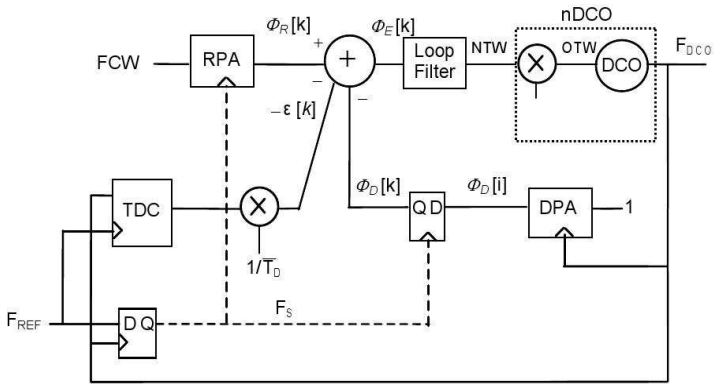


FIG. 5.4 – Synthèse de fréquence RF par ADPLL : principaux blocs fonctionnels et synchronisations.

5.2.1 Simulation

Outre la compréhension du fonctionnement, un premier travail est la *simulation* de l'architecture. Compte tenu des fréquences mises en jeu, cette simulation directe est difficile à envisager car elle conduirait à des temps de simulation considérables et des quantités de données énormes. Pour traiter ce point, et disposer d'un moyen d'ajuster et d'optimiser les différents paramètres de la boucle, nous avons développé un *modèle comportemental complet* du système, qui permet de le simuler avec une charge en calcul très raisonnable.

Les composants limitants sont l'accumulateur de phase du DCO et le TDC (avec une résolution de 10 à 20 ps, qui requièreraient une fréquence de simulation extrêmement élevée. L'astuce du modèle de simulation est le fait que les sorties de ces composants peuvent être exprimées uniquement au rythme Fref (ce qui se comprend comme une suppression de porteuse). L'idée est que la variable intéressante est le décalage de phase entre l'horloge de référence Fref et la sortie du DCO Fdco, et de rechercher à ne travailler que sur ce retard. On a alors pu établir une récurrence permettant d'exprimer directement ce décalage τ en fonction de l'état précédent et des caractéristiques du système au rythme Fref. Ce modèle est décrit dans l'article de congrès [96]. À ce sujet, B. Staszewski nous a écrit : « *I read your paper with great interest. Certain passages looked like something I would have written but could not due to the confidential nature of the design and application. I welcome a new and fresh angle on the stuff that I have been working on for quite a long time. I wish we could co-operate on this research. This topic requires a special set of skill and you have excelled.* »

Cependant, le modèle comportemental n'était valide que pour les modes de calibration et d'acquisition. En mode de poursuite, un modulateur $\Sigma\Delta$ est activé pour introduire un dithering à l'entrée du DCO de sorte à atteindre la résolution voulue. Dans ce cas, on a dû (et pu) modifier les équations précédentes avec l'introduction de deux variables intermédiaires, dont une seule est évaluée à la fréquence de fonctionnement du $\Sigma\Delta$, les autres restant évaluées au rythme Fref. Dans la référence [97], on a également intégré une modélisation plus fine du TDC, prenant en compte les effets de quantification temporelle. Ceci permet de simuler le modèle complet avec une complexité très inférieure à celle, déraisonnable, que l'on obtiendrait par simulation directe. Un dernier point est que l'on cherche, bien sûr, à déterminer la densité spectrale de la forme d'onde en sortie du DCO. Si $f_i(t) = f_0 + \Delta f g(t)$ désigne la fréquence instantannée, $\theta_i(t)$ la phase associée, le signal de sortie s'écrit $s(t) = \cos(\theta_i(t))$. On montre simplement que le spectre correspondant peut s'exprimer directement à partir du signal en bande de base :

$$S(f) = \frac{\Delta f}{2}\left(\frac{G(f-f0)}{(f-f0)} + \frac{G(f+f0)}{(f+f0)}\right)$$

avec $S(f_0) = 1/2 - \pi\Delta f/T \int_0^T \int_0^t g(\tau)d\tau dt$. Ceci permet, à nouveau, de se passer de l'évaluation du signal de sortie, à la fréquence la plus élevée Fdco. À partir de l'ensemble de ces résultats, le modèle comportemental permet une simulation rapide, avec des résultats identiques à ceux délivrés par le modèle

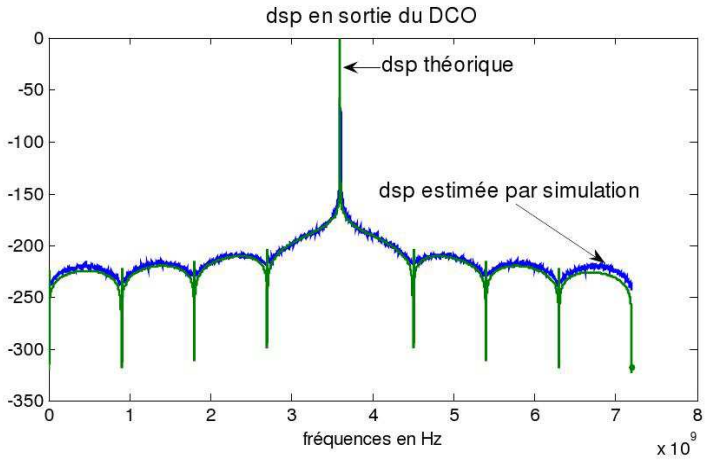


FIG. 5.5 – Comparaison obtenu entre la formule exacte et un résultat de simulation

de simulation "circuits" développé en VHDL, tout en permettant l'accès aux variables d'intérêt.

Une autre contribution, essentiellement due à Geneviève, est présentée en [98], est le calcul théorique de la densité spectrale de puissance de la sortie d'un DCO soumis à un dithering (celui-ci consiste à commuter deux valeurs de capacité à une fréquence élevée, ce qui permet d'améliorer la résolution fréquentielle). En supposant une commutation entre deux capacités suivant une loi binaire non équilibrée, on obtient l'expression exacte de la dsp : l'opération de dithering est assimilée à une modulation de fréquence numérique dans laquelle l'horloge des données est la fréquence de dithering $f_{dither} = 1/T$ et l'excursion de fréquence Δf est égale à la différence entre les deux valeurs de fréquence f0 et f1 correspondant aux deux capacités commutées. La fréquence instantanée est alors un signal de type NRZ à deux états $fc \pm \Delta f/2$. L'indice de modulation correspondant est égal à : $h = \Delta f / f_{dither}$.

En utilisant le principe de la décomposition de Laurent pour les signaux modulés en fréquence à phase continue, on peut récrire le signal généré comme un signal PAM (Pulse Amplitude Modulation), où l'impulsion de mise en forme est une cloche de Laurent. Dès lors, on peut calculer la densité spectrale par la formule de Benett. On retrouve en particulier une constation expérimentale : le fait que la dsp du signal de sortie de l'ADPLL présente une décroissance en $1/f^4$ quel que soit le type de filtre de boucle.

5.2.2 Contributions adaptatives

D'autre part, le système comporte deux calculs d'inversion sur des données qui peuvent être variables dans le temps. Nous avons construit deux algorithmes adaptatifs permettant de suivre ces variations et surtout d'alléger le coût de calcul global. Le premier permet d'éviter une inversion en passant par le biais d'une estimation récursive de l'inverse. Le second permet de calculer la normalisation du gain du DCO.

La figure Fig. 5.4 fait apparaître que le calcul de l'erreur fractionnaire $\epsilon(k)$ demande la multiplication par l'inverse de la période (moyenne) du DCO. En posant $\alpha = 1/T$, et en utilisant le critère quadratique $\mathrm{E}\left[\alpha T - 1\right]$, il est aisé de dériver un algorithme du gradient permettant le calcul de cet inverse. On a aussi proposé au passage de calculer la période moyenne par moyenne exponentielle plutôt que par fenêtre glissante, ce qui permet d'économiser un peu en mémoire. Cette approche a été implantée "en vrai" dans un prototype de PLL numérique pour émetteur FM réalisé et testé par Cyril Joubert chez ST.

Pour introduire une robustesse vis-à-vis des variations du gain du DCO, et une facilité de commande, il est utile d'introduire une normalisation par le gain du DCO. Nous avons proposé également une structure adaptative permettant d'estimer le facteur de normalisation. Simultanément, une formule équivalente a été publiée et évaluée par Staszewski et al. [99]. Notre petit développement a toutefois le mérite d'apporter une justification de l'algorithme ad-hoc proposé. Ces deux algorithmes sont présentés page 176 et suivantes.

5.3 Référence très haute fréquence stabilisée à base de BAW

Les systèmes de communication modernes réclament des oscillateurs de référence à très haute fréquence, notamment pour l'étape de synchronisation entre systèmes. Les nouvelles applications qui requièrent des circuits bas coût, précis et stables peuvent profiter de solutions alternatives aux quartz conventionnels. En effet, un quartz est un composant relativement coûteux et n'est pas intégrable. Bien que les oscillateurs à quartz présentent une très grande pureté spectrale, ils opèrent à une fréquence basse (26 MHz) et requièrent des multiplicateurs de fréquence nombreux. Une alternative intéressante est la possibilité d'utiliser un résonateur à onde acoustique volumique BAW (Bulk Acoustic Wave) dans une structure d'oscillateur. Le facteur de qualité élevé, couplé à la très haute fréquence de résonance des résonateurs BAW permet de réaliser des résonateurs RF présentant une pureté spectrale équivalente à celle des oscillateurs à quartz. Un des intérêts des résonateurs à BAW est leur très faible taille, qui permettra leur intégration avec le transceiver dans un seul "chip" radio. C'est sur cette problématique que Nxp nous a proposé de travailler, dans le cadre du suivi de la thèse en contrat Cifre de Pierre Guillot (Coencadrement C. Berland/J.-F. Bercher).

Cependant, comparé aux oscillateurs à quartz, les oscillateurs à BAW ont une précision de fréquence plus faible, et il convient donc de disposer d'une possibilité d'ajuster la fréquence de l'oscillateur, sur une gamme de fréquence relativement large. À partir d'une préétude sur la dispersion des composants, on a fixé une contrainte d'un ajustement possible sur une bande de 500 kHz, avec un pas de 0.4ppm, soit 800 Hz pour une fréquence d'oscillation cible de 2 GHz.

Par ailleurs, l'imprécision en fréquence dépend de
- l'erreur initiale sur le résonateur BAW (dispersion des caractéristiques nominales),
- le glissement de fréquence avec les variations de température, qui peuvent être particulièrement sévères dans les mobiles,
- variation de la fréquence avec la valeur de la tension d'alimentation,
- phénomènes de vieillissement.

Il convient ainsi de contrôler les différentes dérives possibles (température, alimentation, vieillissement) et stabiliser le fonctionnement, au cours du cycle de vie du mobile, et ce, d'autant plus que les fréquences de fonctionnement sont élevées. Il convient également de concevoir des solutions qui prennent en compte la variabilité statistique des composants et réalisations, et comprennent une dimension fiabilité.

Une première étude sur la réalisation d'une référence à 2GHz, construite autour d'un BAW, est présentée dans [100] (septembre 2008), où l'on décrit le principe et les caractéristiques du système, avec notamment le dimensionnement d'une banque de capacités qui servira à contrôler la fréquence de l'oscillateur. Nous avons choisi d'utiliser deux banques de capacités recouvrantes. La première, grossière, permet un réglage sur 1024 ppm, avec un pas de 32 ppm. La banque fine, quant-à-elle permet un réglage sur 50 ppm avec un pas de 0.4 ppm. Ces deux banques, qui font finalement écho à la solution adoptée pour la partie DCO de l'ADPLL, permettront un réglage et un contrôle numérique de l'oscillateur, pour lesquels il faudra développer les stratégies adéquates.

Dans ce premier article [100], que nous rapportons en annexe, pages 181 à 185, on aborde la faisabilité de principe du dispositif, et les différents résultats sont établis par modélisation et simulation. Depuis, Pierre a fait réaliser la référence de fréquence en technologie CMOS 65nm. Une photographie est présentée Fig. 5.6.

Cette réalisation a permis de mesurer pratiquement les performances obtenues et de les confronter aux différents modèles de simulation. En pratique, on a obtenu une grande pureté spectrale (-128dBc/Hz à 100 kHz de la fréquence porteuse). On a noté un décalage de la fréquence d'oscillation. Une faible consommation (0.9 mW) a été mesurée. Sur cette réalisation et les mesures, où ma contribution est essentiellement d'ordre méthodologique, nous avons écrit un petit article de revue [101], accepté très récemment, pour présenter la réalisation, les résultats de mesure et leur comparaison à ceux obtenus par les modèles numériques. Une communication en congrès reprend également une partie de ces résultats [102].

Ainsi, les objectifs en terme de pureté et de consommation ont été atteints et on évalue maintenant

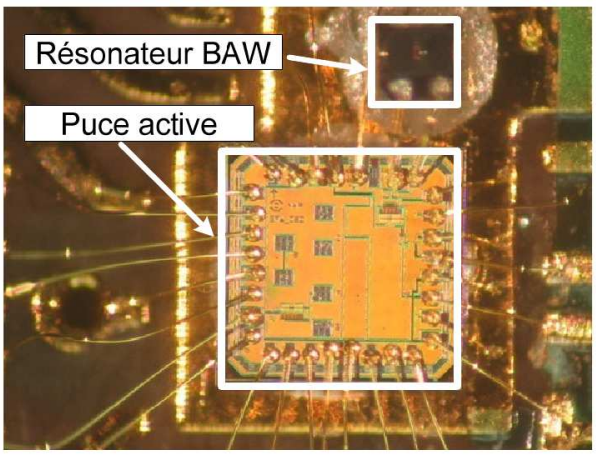


FIG. 5.6 – Réalisation : Résonateur BAW et partie active

l'aspect calibration et contrôle. Un modèle physique du système a été réalisé, notamment concernant la variation de la fréquence de sortie avec les capacités. Il a été complété par deux modèles des dérives en fonction de la température et de la tension d'alimentation. Pour cela, nous avons simplement ajusté les paramètres de polynômes du second degré à l'aide de mesures. Au total, on obtient un modèle direct dans lequel la fréquence de sortie dépend des conditions de réalisation, de la température, de la tension d'alimentation et des valeurs prises par les capacités de contrôle (toutes ces grandeurs sont connues ou mesurables). La caractérisation du système impose d'estimer les différents paramètres, à savoir les gains des banques de capacités et les coefficients des polynômes, et de poursuivre leur évolution éventuelle.

Pour les étapes d'estimation et de contrôle, nous développons une solution fondée sur une modélisation d'état et un filtrage de Kalman. Le problème est difficile car il est très sous déterminé (seule la fréquence de sortie est observable), mais on a une bonne idée de l'état moyen qui servira de condition initiale et des incertitudes associées qui permettront de définir la matrice de covariance a priori. Le filtre de Kalman permet alors d'ajuster les paramètres du modèle d'état et de contrôler la précision des estimées (via les matrices de covariance a posteriori). Deux modes de fonctionnement seront utilisés : une étape de calibration initiale, où l'on disposera d'une fréquence de référence externe, et une possibilité de poursuite des paramètres, en profitant de la fréquence de référence délivrée par le réseau. Pour ce faire, un dispositif de mesure de l'écart en fréquence a été développé. La partie traitement et contrôle sera développée en VHDL et implantée sur FPGA.

Chapitre 6

Synchronisation des voies dans les architectures polaires

Dans le chapitre précédent, lors de l'exposé de la nouvelle architecture d'émetteur, nous avons déjà abordé le principe des architectures EER, dites aussi polaires, qui consistent à séparer les traitements sur l'enveloppe et la phase du signal de communication.

Dans une architecture polaire, le signal d'enveloppe et le signal de phase subissant des traitements différents, une désynchronisation entre l'enveloppe et la phase peut apparaître. Cette désynchronisation est liée à la différence de temps de traitement entre les deux voies, au retard de groupe des filtres, tout autant que les petites contributions de tous les petits éléments analogiques. Cette désynchronisation entraîne une remontée du plancher de bruit et une augmentation de l'EVM (Error Vector Magnitude). Ce problème, qui n'était pas réellement considéré jusqu'à ces dernières années, devient maintenant d'autant plus crucial que la période symbole diminue, avec l'augmentation des débits [103, 104]. Cela devient un sujet de préoccupation qui est évoqué dans plusieurs publications, par exemple [105, 106, 107]. En production, même avec la connaissance globale des retards apportés par les différents éléments, les retards doivent être ajustés au cas par cas pour tenir compte de la dispersion des éléments. Par ailleurs, les caractéristiques analogiques sont affectées par des dérives dans le temps, par une variation avec la température, de la tension d'alimentation, et par un effet de vieillissement. Enfin, les solutions sont de plus en plus multimodes et les caractéristiques peuvent varier rapidement lorsque l'on passe d'un mode à un autre. On pourrait imaginer que le passage à des architectures numériques efface ce problème. En fait il n'en est rien, comme en témoigne la publication récente [108] (ISCAS - mai 2008) de K. Waheed, R. B. Staszewski and S. Rezeq intitulée de manière évocatrice « *Curse of Digital Polar Transmission : Precise Delay Alignment in Amplitude and Phase Modulation Paths* ». Dans ce travail, les auteurs montrent que l'assymétrie du traitement, l'indépendance des horloges sous-jacentes et les interfaces analogiques imposent, pour satisfaire aux contraintes strictes des standardisations, un alignement extrêmement précis des deux voies. On peut noter que les courbes sur la dégradation des performances fournies dans ce papier confirment celles que nous avons données dans [109].

Dans un premier paragraphe, nous examinons l'effet d'un mauvais alignement sur les performances du système, en introduisant un modèle de simulation par technique de Monte-Carlo [110]. Dans la section 6.2, nous discutons de la faisabilité de cet alignement, avant de décrire rapidement, en section 6.3, un algorithme de correction de ces défauts [111, 109]. Enfin, dans la section 6.4, nous proposons une version sous-optimale de cet algorithme, ainsi que des extensions permettant de prendre en compte des méconnaissances du gain et de la phase en sortie de l'émetteur [112].

6.1 Effet d'un mauvais alignement

Pour des modulations à grand nombre d'états, ou à multiporteuses, le signal émis $x(t)$ peut souvent être modélisé comme gaussien complexe (et c'est bien d'ailleurs ce qui pose des problèmes car les grandes

variations d'amplitude entraînent des difficultés et des distorsions pour tous les éléments non-linéaires). Dans le cas d'un processus complexe circulaire (possédant donc un spectre passe bande symétrique autour de sa fréquence centrale et dont les composantes I et Q sont distribuées avec une loi gaussienne), il est bien connu que l'enveloppe et la phase, au même instant, sont indépendantes et distribuées selon des lois respectives Rayleigh et uniforme. Ce résultat est également vrai à deux instants différents, comme nous l'avons montré dans [113]. Ceci signifie, en clair, que pour un processus gaussien, l'enveloppe et la phase sont toujours indépendantes, sans aucune référence au coefficient de corrélation, et ce quelque soit le décalage temporel entre les deux composantes. Ainsi, les statistiques à un instant ne portent aucune information sur le décalage entre l'enveloppe et la phase. Il faudra donc au moins s'intéresser aux statistiques d'ordre 2, et c'est justement celles qui nous importent, puisqu'une des conséquences importantes d'un mauvais alignement est une remontée du plancher de bruit.

Dans un premier temps, j'ai cherché à étudier les propriétés du processus

$$x_\Delta(t) = \rho(t)\cos(\phi(t-\Delta)),$$

où ρ et ϕ sont l'enveloppe et la phase issus d'un processus gaussien connu. En effet, aussi bien pour la détermination analytique de la densité spectrale $S_{x_\Delta x_\Delta}(f)$ que pour les différentes propriétés des algorithmes du second ordre que l'on peut avoir à mettre en œuvre, il est utile de caractériser le signal précédent, au moins au second ordre. En fait, on voit que pour déterminer la fonction de corrélation $R_{x_\Delta x_\Delta}(\tau)$ associée à $x_\Delta(t)$, il faut utiliser la loi du processus gaussien aux quatre instants $\{t, t-\tau, t-\Delta, t-\tau-\Delta\}$, et l'intégrer par rapport à 4 des 8 variables après passage en coordonnées polaires. On s'aperçoit, dans mon cas après beaucoup d'efforts, mais avec la lecture fascinante de l'ouvrage de Levine [1], qu'il n'y a pas d'expression utilisable. Tout juste existe-t-il des développements en série (où les coefficients sont des intégrales de fonctions de Bessel) pour la fonction d'autocorrélation du module et pour celle de la phase. Notons qu'il est aisé de vérifier que même dans le cas $\Delta = 0$, on n'a pas $p(\rho(t), \phi(t), \rho(t-\tau), \phi(t-\tau)) = p(\rho(t), \rho(t-\tau))p(\phi(t), \phi(t-\tau))$, ce qui montre que les processus d'enveloppe et de phase ne sont pas indépendants.

Avec Antoine Diet, dont j'ai dirigé les travaux sur cette partie de sa thèse de doctorat, nous avons cherché à évaluer simplement, forcément numériquement, l'effet du décalage entre les deux voies sur la densité spectrale de sortie. Bien sûr il est possible de simuler une chaîne d'émission complète et d'évaluer, par méthode de Monte-Carlo, les performances spectrales obtenues en faisant varier tel ou tel défaut. Un premier inconvénient est que les résultats peuvent dépendre d'autres caractéristiques ou défauts du modèle de simulation. Un second inconvénient est la lourdeur du processus et le temps de simulation associé. Dans ce travail, nous avons simplement choisi de générer des vecteurs aléatoires gaussiens complexes aux 4 instants voulus, avec une structure de corrélation est connue (c'est la carré de convolution de la réponse impulsionnelle du filtre d'émission) puis de calculer la fonction de corrélation empirique associée :

$$\hat{R}_{x_\Delta x_\Delta}(\tau) = \frac{1}{K}\sum_{i=1}^{K} \rho_1(i)\phi_2(i)\rho_3(i)\phi_4(i),$$

où les ρ_k et ϕ_k, $k = 1 \ldots 4$, sont les variables aléatoires d'enveloppe et de phase aux 4 instants. Nous avons pu également calculer la variance d'estimation, qui décroît évidemment en $1/K$, ce qui a permis d'affecter des intervalles de confiance aux résultats. Ce travail a été publié en [110], et présenté de manière plus détaillée dans le manuscrit de thèse d'Antoine Diet (Partie V). La figure Fig. 6.1 montre le comportement de la densité spectrale de puissance dans le cas d'un décalage entre voies dans le cas de l'Hiperlan2. On constate que les spécifications spectrales ne sont plus respectées dès que le décalage est supérieur ou égal à 2 ns, ce qui représente 10% de la période symbole. Il apparaît donc impératif, surtout avec la montée en fréquence, de chercher à corriger les défauts d'alignement.

6.2 Faisabilité de l'alignement entre voies

Discutons encore un instant de la faisabilité de l'alignement. Comme on l'a vu, on doit s'intéresser à une statistique à deux instants, au moins, sur la sortie. Si $z(t)$ est la sortie réelle, où l'enveloppe a subi un

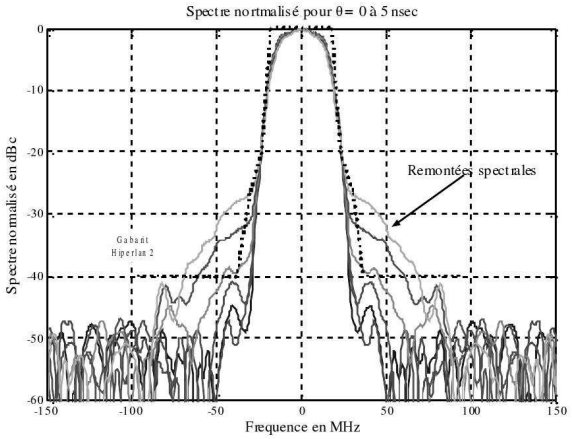


FIG. 6.1 – Remontées spectrales dans le cas d'un décalage enveloppe-phase, de 0 à 5ns

décalage de Δ_1 et la phase un décalage de Δ_2, la fonction d'autocorrélation s'écrit

$$R_{zz}(\tau) = \mathrm{E}\left\{\rho(t-\Delta_1)\cos(\phi(t-\Delta_2))\rho(t-\tau-\Delta_1)\cos(\phi(t-\tau-\Delta_2))\right\}, \tag{6.1}$$

dont, on l'a déjà indiqué, on n'a pas d'expression explicite. Il est donc difficile d'imaginer une procédure qui reposerait sur cette seule fonction d'autocorrélation. Cependant, si on introduit un retard μ sur la voie enveloppe, par exemple, en notant $\Delta = \Delta_2 - \Delta_1$ le retard relatif et en utilisant la stationnarité du signal,

$$R_{zz}(\tau,\mu) = \mathrm{E}\left\{\rho(t-\mu)\cos(\phi(t-\Delta))\rho(t-\mu-\tau)\cos(\phi(t-\Delta-\tau))\right\}. \tag{6.2}$$

Évidemment cette nouvelle fonction de corrélation coïncidera avec la fonction de corrélation $R_{xx}(\tau)$ du signal initial, aligné, quand $\mu = \Delta$. Dès lors, une procédure possible consiste à rechercher une valeur de μ qui minimise une distance entre $R_{zz}(\tau,\mu)$ et $R_{xx}(\tau)$, pour tout τ, par exemple

$$\int w(\tau)^2 \left(R_{zz}(\tau,\mu) - R_{xx}(\tau)\right)^2 \mathrm{d}\tau, \tag{6.3}$$

ou $w(\tau)^2$ est une fonction de fondération, ou, de manière équivalente, dans le domaine fréquentiel,

$$\int |W(f)|^2 \left(S_{zz}(f,\mu) - S_{xx}(\tau)\right)^2 \mathrm{d}f. \tag{6.4}$$

À partir de ces idées, on note que l'on a évidemment besoin d'un retour sur la sortie, c'est-à-dire d'une manière d'acquérir ou de calculer une quantité caractéristique à partir des données de sortie, et de comparer celle-ci à une statistique connue, ou estimée, de l'entrée (du signal aligné). On devra donc implanter une boucle de retour, qui mettra en œuvre une démodulation et numérisation d'une image du signal en sortie de l'émetteur. Notons qu'une telle boucle de commande est de toute façon souvent déjà présente, pour des problèmes de calibration et contrôle (linéarisation de l'amplificateur de puissance par exemple).

En réalité, ces deux possibilités sont difficiles à mettre en œuvre pratiquement, car elles nécessitent l'estimation de sommes de fonction de corrélations, ou de spectres, (et sans doute de leurs dérivées) ce qui est lourd et statistiquement délicat. De plus l'algorithme de contrôle ne peut pas être caractérisé précisément, s'il peut être établi, du fait du manque d'expression explicite pour la corrélation.

6.3 Un algorithme de correction

La procédure que nous avons proposée, cf Fig. 6.2, consiste à comparer directement la sortie au signal numérique initial, bien sûr après démodulation et numérisation (suivant les cas par sous-échantillonnage, par démodulation simple ou en quadrature). Cette comparaison est effectuée en moyenne quadratique, et deux retards μ_1, μ_2 sont ajustés de sorte à identifier et compenser les retards inconnus Δ_1 et Δ_2 introduits par le système.

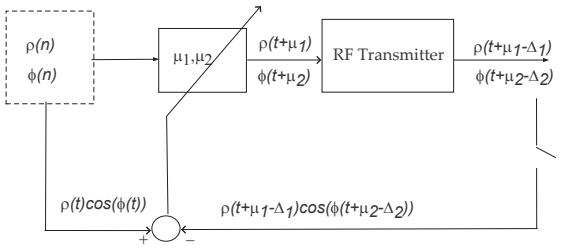


FIG. 6.2 – Principe de correction des retards.

Après l'introduction des deux avances μ_1 et μ_2, la sortie du système s'écrit $z(t) = \rho(t+\mu_1-\Delta_1)\cos(\phi(t+\mu_2-\Delta_2))$. La distance quadratique au signal $x(t)$ fournit alors le critère

$$J(\mu_1,\mu_2) = \mathrm{E}\big[|x(t) - \rho(t+\mu_1-\Delta_1)\cos(\phi(t+\mu_2-\Delta_2))|^2\big],$$

En tenant compte de l'indépendance entre l'enveloppe et la phase à un instant, le critère se réduit simplement à

$$J(\mu_1,\mu_2) = 4R(0,0) - 4R(\mu_1-\Delta_1,\mu_2-\Delta_2) \tag{6.5}$$

où $R(\tau_1,\tau_2) = \mathrm{E}\left[\rho(t)\cos(\phi(t))\rho(t-\tau_1)\cos(\phi(t-\tau_2))\right]$ est une sorte de fonction de corrélation. On pourra noter que par l'inégalité de Schwarz $R(0,0) \geq |R(\tau_1,\tau_2)|$. La figure Fig. 6.3 donne l'allure du critère, obtenue par l'évaluation numérique de (6.5). On observe un minimum global clair, qui pourra être atteint par une méthode de descente, pour des retards inférieurs à une période symbole ; au delà, des minima secondaires apparaissent.

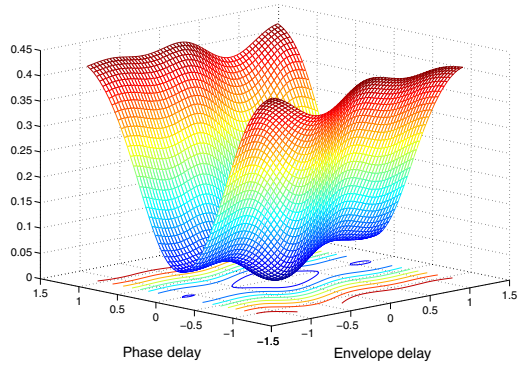


FIG. 6.3 – Critère quadratique entrée-sortie en fonction des décalages Δ_1 et Δ_2 sur les deux voies.

Nous avons proposé un premier algorithme de correction, reposant sur ce critère, dans [111]. On propose simplement de minimiser le critère à l'aide d'un algorithme de de gradient stochastique, ce qui conduit à des itérations du type

$$\begin{aligned} \mu_1(n+1) &= \mu_1(n) + \gamma_1(n)\left.\frac{d\rho(u)}{du}\right|_{t_1(n)} \cos\left(\phi(t_2(n))\right) e(t) \\ \mu_2(n+1) &= \mu_2(n) + \gamma_2(n)\rho(t_1(n))\left.\frac{d\cos\phi(u)}{du}\right|_{t_2(n)} e(t) \end{aligned} \tag{6.6}$$

avec $t_1(n) = nT_s + \mu_1(n) - \Delta_1$ et $t_2(n) = nT_s + \mu_2(n) - \Delta_2$, et $e(t) = x(t) - z(t)$.
À la suite de cette communication, nous avons été invités à compléter ce travail et soumettre une version étendue sous forme d'un article de revue [109], qui est proposé ici en annexe, pages 190 à 203. Bien

sûr, comme dans les algorithmes de synchronisation intervenant au niveau du récepteur, on doit disposer, et reconstituer au final des valeurs entre les échantillons disponibles : la procédure doit alors employer soit un taux de suréchantillonnage complètement déraisonnable, soit des dispositifs d'interpolation numérique. Le rôle, le choix et les conséquences de l'interpolateur a été étudié avec quelques détails dans l'article en question. On a également étudié un modèle approché de notre problème initial, et simplement considéré $z(t) = x(t - \Delta)$, une version retardée d'un signal original $x(t)$. Afin d'identifier et corriger le retard, on utilise

$$\mu(n+1) = \mu(n) + \gamma \left.\frac{dx(u)}{du}\right|_{u=t-\Delta+\mu(n)} [x(t) - \tilde{x}(t - \Delta + \mu(n))] \tag{6.7}$$

avec $\tilde{x}(t - \Delta + \mu(n))$ la valeur interpolée au temps $(t - \Delta + \mu(n))$. En utilisant ce modèle simplifié, on a pu établir, et on l'a vérifié numériquement sur le problème « complet », que

– le temps de convergence décroît exponentiellement avec le pas d'adaptation,
– un biais d'estimation existe, lié à l'ordre de l'interpolateur et au fait que l'erreur quadratique totale est minimum avec un peu de biais,
– la variance est une fonction (approximativement) linéaire du pas d'adaptation,
– l'EVM s'accroît linéairement avec le pas d'adaptation.

On a aussi montré que l'algorithme proposé permet d'augmenter considérablement les performances, que ce soit en terme spectral ou d'EVM. Les figures Fig. 6.4 et Fig. 6.5 en donnent une illustration.

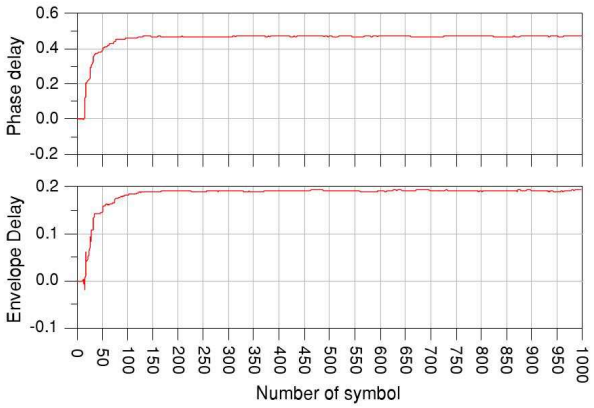


FIG. 6.4 – Convergence de μ_1 et μ_2 vers les vrai retards $\Delta_1 = 0.47T_s$ et $\Delta_2 = 0.18T_s$.

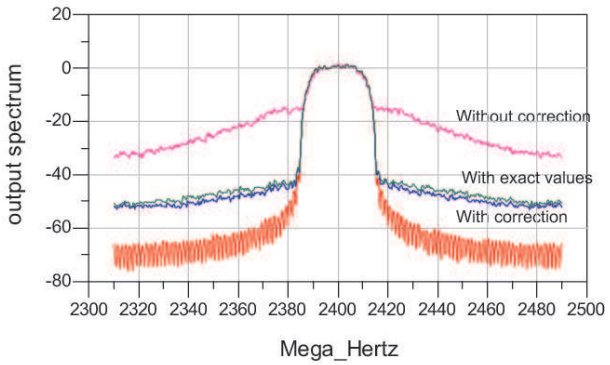


FIG. 6.5 – Comparison des spectres idéal, non corrigé et corrigé, dans le cas de retards $\Delta_1 = 0.47T_s$ et $\Delta_2 = 0.18T_s$.

6.4 Extensions

Un défaut de la procédure précédente est que le terme d'adaptation de l'algorithme est calculé sur le signal de sortie et nécessite de disposer séparément de l'enveloppe et de la phase de la sortie. Ceci impose nécessairement de procéder à une démodulation en quadrature, puis à une conversion en représentation polaire. Dans la communication [113], nous avons considéré une procédure d'identification des retards, puis de correction a posteriori. Dans ce cas, le terme correctif dépend uniquement de l'entrée, qui par ailleurs est disponible en module et phase (on a tout fait pour). Par conséquent, on peut alors se contenter d'une démodulation simple, ou même d'une conversion directe par sous-échantillonage, ce qui allège beaucoup le coût d'implantation.

C'est à partir de ces observations que nous avons ensuite développé une version sous-optimale de l'algorithme de correction, qui permet elle-aussi d'éviter une démodulation en quadrature dans la boucle de retour. L'idée est simplement se substituer aux quantités qui nécessiteraient la séparation de l'enveloppe et de la phase, en sortie, les mêmes quantités exprimées aux instants d'entrée (i.e. sans prendre en compte le décalage temporel) : il est clair que lorsque l'algorithme de correction approche la solution correcte $\mu_i \rightarrow \Delta_i$, on a bien sûr $t_i = t + \mu_i - \Delta_i \rightarrow t$, et on peut approximer, par exemple

$$\left.\frac{d\rho(u)}{du}\right|_{t_1} \cos\phi(t_2) \approx \left.\frac{d\rho(u)}{du}\right|_{t} \cos\phi(t)\,. \tag{6.8}$$

Ceci entraîne une réduction du domaine de convergence, qui a été décrit, mais la procédure reste fonctionnelle et performante. En pratique, la sortie de l'émetteur est entachée d'un gain et d'une rotation de phase, ce qui n'a pas beaucoup d'importance habituellement, compte tenu du fait que le signal va subir d'autres avanies lors de la transmission et que le récepteur est équipé pour compenser ces défauts. Par contre, dans le cas qui nous occupe, cela a certains effets puisque l'on compare le signal de sortie à celui d'entrée. On montre analytiquement que l'argument du minimum du critère est inchangé, c'est-à-dire que la procédure de correction est robuste à ces défauts. Cependant, l'erreur quadratique à l'optimum n'est plus nulle, ce qui entraîne une augmentation de la variance sur les estimées des retards. Nous avons alors développé une extension de la procédure, par l'introduction d'une étape d'identification du gain et de la phase dans la boucle de retour, qui permet de prendre en compte et compenser ces méconnaissances. Les résultats de simulation montrent une nette augmentation des performances. Nous avons aussi étudié le comportement de cette solution en contexte non stationnaire. Enfin, le problème a été simulé, et les performances mesurées, en prenant en compte les effets de quantification. Pour le détail de ces résultats, les lecteurs pourront se reporter en pages 208 à 215 où est reproduit un article actuellement soumis à *IEEE trans. on Circuits and Systems I* [112].

Nous avons d'ores et déjà prolongé ces résultats, analytiquement et en simulation, à l'architecture LINC, potentiellement très performante, mais qui confrontée aux mêmes types de désynchronisation, peut-être de manière plus aigüe encore. Rappelons très rapidement que l'idée du LINC est d'écrire le signal de départ comme la somme de deux signaux à enveloppe constante, selon

$$X(t) = \rho(t)e^{j\phi(t)} = R\cos\left(\theta(t)\right)e^{j\phi(t)} = R\left(e^{j(\phi(t)+\theta(t))} + e^{j(\phi(t)-\theta(t))}\right)$$

puis d'amplifier séparément les deux signaux avant de les recombiner au niveau de l'antenne.

Du fait que l'architecture comporte une opération très non linéaire, le SCS (Signal Components Separation), il est délicat d'effectuer une correction directe et notre démarche est une démarche d'identification des retards puis de correction a posteriori. Les résultats obtenus sont excellents. Une autre perspective consistera a également prendre en charge des ondulations résiduelles du gain dans la bande passante (ce qui revient en fait à remplacer la correction du retard par un égaliseur, mais au niveau de l'émetteur).

Par ces travaux, on confirme que la conception et l'optimisation des émetteurs numériques peut ainsi bénéficier de l'apport d'outils de traitement du signal pour l'optimisation de la "dirty RF" : c'est le point de vue que nous avons présenté dans [114] (Conférence *invitée* - C. Berland).

Chapitre 7

Points de suspension

Comme je l'écrivais en préambule, la présentation de ce mémoire n'est ni une finalité ni un achèvement mais simplement un signe de ponctuation dans une histoire que je souhaite plus longue. J'ai montré dans les pages précédentes que j'ai développé des recherches suivant deux sillons parallèles dans ma géométrie personnelle. Le premier correspond à mes intérêts historiques en théorie et traitement de l'information, avec en particulier l'étude et la caractérisation d'entropies et autres mesures d'information, et leurs applications à divers problèmes. Le second axe concerne les apports du traitement du signal pour la conception et l'évaluation des émetteurs récepteurs de communication. Des perspectives naturelles se dégagent sur les deux axes.

Comme j'ai débuté le mémoire par la partie entropies, je terminerai, en forme de boucle, en envisageant d'abord les perspectives en radiotechnique.

L'un des points de départ de ces travaux menés avec C. Berland est la recherche d'une structure d'émetteur universel, capable de prendre en charge des signaux de nature différente. La première idée étant ainsi de s'intéresser aux structures décomposant les signaux en deux composantes d'enveloppe et de phase. C'est sur cette structure que l'on a proposé des évolutions et des dispositifs de correction afin d'en améliorer les performances. Il est possible de développer d'autres méthodes de compensation des défauts. Une idée fort simple, et qu'il faudra évaluer prochainement, analytiquement et en pratique, est de chercher à minimiser la puissance, et donc les remontées spectrales dans une bande d'intérêt. Une autre perspective consistera, en plus des corrections de retards, a également prendre en charge des ondulations résiduelles du gain dans la bande passante, ce qui revient en fait à intégrer un égaliseur, qu'il faudra imaginer léger, au niveau de l'émetteur. Une autre prolongation immédiate est de considérer l'architecture LINC qui consiste elle-aussi à décomposer le problème difficile en deux problèmes plus simples, mais qui souffre également de problèmes de désynchronisation. Il est tout-à-fait envisageable que les structures d'algorithme de correction que nous avons étudiées puissent apporter ici des résultats intéressants.

En poursuivant sur cette idée de simplifier la réalisation, il serait également intéressant de rechercher d'autres types de représentation des signaux, par décompositions sur des ensembles de fonctions, ou par des dispositifs de codage simple qui permettraient d'envisager d'autres structures de calibration au niveau de l'émetteur. Le cheminement des signaux comprend des non-linéarités, notamment l'étage de puissance, auxquelles je souhaite m'intéresser, tout en étant bien conscient que le sujet a été largement exploré, mais tout le monde n'a pas la chance de pouvoir étudier [1]. Toujours au niveau de l'émetteur, les travaux menés avec P. Guillot concernant la réalisation d'une fréquence de référence à très haute fréquence nous conduisent à l'étude d'un problème direct, et au développement de solutions d'estimation et de contrôle à base de modélisation d'état et de filtrage de Kalman. À trop promouvoir l'intégration du traitement du signal tôt dans les traitements, on se retrouve à chercher à numériser au plus tôt, éventuellement avec des défauts qui devront être compensés par la suite. C'est sur cette problématique que débute un autre travail avec A. Lesellier, sur laquelle on envisage des solutions à base de bancs de filtres hybrides analogique/numérique, avec des contraintes sur la réalisabilité des solutions. D'autres possibilités, comme l'utilisation de résultats sur l'échantillonnage non-uniforme demandent à être explorées.

Du côté des entropies, on a des possibilités immédiates de prolongation des travaux exposés. Il ne serait pas difficile par exemple d'exploiter les fonctionnelles entropiques associées à la divergence de Rényi dans le cadre des problèmes inverses. Sur ce sujet, il serait intéressant d'imaginer des algorithmes de relaxation sur le paramètre α. Sur ce thème, un autre travail à reprendre est celui qui concerne la technique de maximum d'entropie itérative exposée en §2.2. On a noté, dans le paragraphe §3.4 que pour les distributions de Tsallis que nous avons étudiées, la transformation paramètre-moyenne n'est pas bijective, ce qui peut mener à des interprétations sympathiques de type transition de phase. Une étude précise sur la caractérisation du phénomène est à compléter. Le rôle de la contrainte de log-vraisemblance moyen, que nous avons utilisé dans [29], cf. §3.2, notamment le lien avec le codage, est encore un point à explorer.

Dans les constructions que nous avons présentées apparait souvent l'"escort distribution" associée à la distribution initiale, et il est possible que celle-ci joue un rôle fondamental dans ces constructions. J'envisage ainsi de me pencher plus précisément sur cette correspondance entre distributions et sur les propriétés de ces escort distribution, qui peuvent être éventuellement à relier aux notions de distributions incomplètes utilisées par Rényi, cf l'excellent livre de Pierre Hammad [115]. Il est possible par exemple de s'intéresser à l'information de Fisher attachée aux escorts, qui semble présenter quelques caractéristiques intéressantes. On pourra noter avec intérêt que la question de la justification des escort distributions fait partie des quelques questions ouvertes relevées dans [38]. L'existence d'une géométrie particulière sous jacente à la famille des exponentielles déformées, à l'image de la géométrie des familles exponentielles, est également un sujet d'intérêt. La divergence de Bregman déformée que l'on a obtenue est un premier élément. Je souhaite aussi continuer à explorer les domaines d'apparition des entropies et distributions de Rényi-Tsallis et en extraire les raisons ou motivations sous-jacentes. Il serait ainsi intéressant d'explorer les connections avec les théories des multifractales. Les lois d'échelle, et sous une certaine forme les lois de Tsallis, apparaissent également dans certains processus multiplicatifs.

Bibliographie

[1] B. R. Levin, *Fondements théoriques de la radiotechnique statistique*. Moscow : Éditions Mir, 1979, vol. 1.

[2] L. Bercher, *L'homœpathie sans mystère*. Paris : G. Doin & Cie, 1937.

[3] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, pp. 620–630, May 1957.

[4] E. T. Jaynes, "Information theory and statistical mechanics. II," *Physical Review*, vol. 108, no. 2, pp. 171–190, Oct. 1957.

[5] G. L. Besnerais, J.-F. Bercher, and G. Demoment, "A new look at entropy for solving linear inverse problems," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1565–1578, 1999.

[6] P. Holejšovská and J.-F. Bercher, "Elimination of the measurement system impact on the measured data by maximum entropy on the mean method," Plzen, Sep. 2003, pp. 192–195.

[7] P. Holejšovská, Z. Peroutka, J. Cengery, and J.-F. Bercher, "Continuous Non-Invasive blood pressure measuring based on the reconstruction by maximum entropy on the mean method," in *17-th Biennial International Eurasip Conference Proceedings*. Brno : Jirí Jan, Jirí-Kozumplík, and Ivo Provazník, Jun. 2004, pp. 121–123.

[8] J.-F. Bercher, "Maximum entropy with fluctuating constraints - The example of K-distributions," *Physics Letters A*, vol. 372, no. 24, pp. 4361–4363, Jun. 2008. PDF

[9] J.-F. Bercher and C. Vignat, "Estimating the entropy of a signal with applications," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1687–1694, 2000.

[10] J.-F. Bercher and C. Vignat, "A Rényi entropy convolution inequality with application," in *EUSIPCO' 2002, XI European Signal Processing Conference*, Toulouse, France., sep 2002, pp. 111–114.

[11] H. Gzyl, "Ill-posed linear inverse problems and maximum entropy in the mean," *Acta Científica Venezolana*, vol. 53, no. 2, pp. 74–93, 2002.

[12] F. Gamboa and E. Gassiat, "Bayesian methods and maximum entropy for ill-posed inverse problems," *The Annals of Statistics*, pp. 328–350, 1997.

[13] P. Holejšovská, Z. Peroutka, and J.-F. Bercher, "Adaptive signal processing using maximum entropy on the mean method and monte carlo analysis," in *MATLAB 2003*, 2003, pp. 191–196.

[14] G. Wilk and Z. Wlodarczyk, "Interpretation of the nonextensitivity parameter q in some applications of Tsallis statistics and Lévy distributions," *Phys.Rev.Lett.*, vol. 84, p. 2770, Aug. 2000.

[15] C. Beck and E. G. D. Cohen, "Superstatistics," *Physica A*, vol. 322, pp. 267–275, May 2003.

[16] F. Sattin, "Non-Gaussian probability distribution functions from maximum-entropy-principle considerations," *Physical Review E*, vol. 68, no. 3, p. 32102, 2003.

[17] H. Touchette and C. Beck, "Asymptotics of superstatistics," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 71, no. 1, pp. 016 131–6, 2005.

[18] B. A. Maguire, E. S. Pearson, and H. A. Wynn, "The time intervals between industrial accidents," *Biometrika*, vol. 39, pp. 168–180, 1952.

[19] J. Beirlant, E. J. Dudewicz, L. Gyorfi, and E. C. van der Meulen, "Nonparametric entropy estimation : An overview," *International Journal of Mathematical and Statistical Sciences*, vol. 6, pp. 17–39, 1997.

[20] J.-F. Bercher and C. Vignat, "Estimating the entropy of a signal with applications," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999. ICASSP '99*, vol. 3, 1999, pp. 1705–1708 vol.3.

[21] C. Vignat and J.-F. Bercher, "Un estimateur récursif de l'entropie," in *17e Colloque sur le traitement du signal et des images*. Vannes : GRETSI,, 1999, pp. 701–704. PDF

[22] K. E. Hild, D. Erdogmus, and J. Principe, "Blind deconvolution with Rényi's minimum entropy," in *Proceedings EUSIPCO 2002*, 2002, pp. 557–560.

[23] D. Erdogmus, K. E. Hild, J. C. Principe, M. Lazaro, and I. Santamaria, "Adaptive blind deconvolution of linear channels using Rényi's entropy with parzen window estimation," *IEEE Transactions on Signal Processing*, vol. 52, no. 6, pp. 1489–1498, 2004.

[24] D. Han, S. Kim, and J. Principe, "Sparse channel estimation with regularization method using convolution inequality for entropy," in *Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on*, vol. 4, 2005, pp. 2359–2362 vol. 4.

[25] S. Cruces and I. Duran, "The minimum support criterion for blind signal extraction : A limiting case of the strengthened Young's inequality," *Lecture notes in computer science*, pp. 57–64, 2004.

[26] S. Cruces, A. Sarmiento, and I. Duran, "The complex version of the minimum support criterion," *Lecture Notes in Computer Science*, vol. 4666, pp. 17–24, 2007.

[27] D. T. Pham, F. Vrins, and M. Verleysen, "On the risk of using Rényi's entropy for blind source separation," *IEEE Transactions on Signal Processing*, vol. 56, no. 10 Part 1, pp. 4611–4620, 2008.

[28] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics*, vol. 46, no. 5, pp. 323–351, 2005.

[29] J.-F. Bercher, "An amended MaxEnt formulation for deriving Tsallis factors, and associated issues," in *AIP Conference Proceedings - Twenty sixth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, A. Mohammad-Djafari, Ed., vol. 872, Nov. 2006, pp. 441–448. PDF

[30] J.-F. Bercher, "Tsallis distribution as a standard maximum entropy solution with 'tail' constraint," *Physics Letters A*, vol. 372, no. 35, pp. 5657–5659, Aug. 2008. PDF

[31] J.-F. Bercher and C. Vignat, "A new look at q-exponential distributions via excess statistics," *Physica A*, vol. 387, no. 22, pp. 5422–5432, Sep. 2008. PDF

[32] J. F. Bercher and C. Vignat, "An entropic view of Pickands' theorem," in *International Symposium on Information Theory ISIT2008*, Toronto, Jul. 2008. PDF

[33] J.-F. Bercher, "On some entropy functionals derived from Rényi information divergence," *Information Sciences*, vol. 178, pp. 2489–2506, Jun. 2008. PDF

[34] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I.* Berkeley, Calif. : Univ. California Press, 1961, pp. 547–561.

[35] L. L. Campbell, "A coding theorem and Rényi's entropy," *Information and Control*, vol. 8, no. 4, pp. 423–429, 1965.

[36] J. Bercher, "Source coding with escort distributions and rényi entropy bounds," *Physics Letters A*, vol. 373, no. 36, pp. 3235–3238, Aug. 2009. PDF

[37] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1, pp. 479–487, Jul. 1988.

[38] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics*. Springer, Apr. 2009.

[39] M. Grendar and M. Grendar, "Maximum probability and maximum entropy methods : bayesian interpretation," in *AIP Conference Proceedings*, vol. 707. IOP Institute of Physics publishing ltd, 2004, pp. 490–494.

[40] J.-F. Bercher, "Possible rationales for Rényi-Tsallis entropy maximization," in *International Workshop on Applied Probabilities (IWAP2008)*, Compiègne, Jul. 2008. PDF

[41] S. Kullback, *Information Theory and Statistics*. Wiley, New York, 1959.

[42] M.-P. Schützenberger, *Contributions aux applications statistiques de la théorie de l'information, thèse de doctorat*. Institut de Statistique de l'Université de Paris, 1954, vol. 3.

[43] A. Wald, "Sequential tests of statistical hypotheses," *Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, Jun. 1945.

[44] I. Csiszar, "Generalized cutoff rates and Rényi's information measures," *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 26–34, 1995.

[45] B. LaCour, "Statistical characterization of active sonar reverberation using extreme value theory," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 310–316, 2004.

[46] J. Pickands, "Statistical inference using extreme order statistics," *The Annals of Statistics*, vol. 3, pp. 119–131, 1975.

[47] A. A. Balkema and L. de Haan, "Residual life time at great age," *The Annals of Probability*, vol. 2, pp. 792–804, Oct. 1974.

[48] N. Markovich, *Nonparametric Analysis of Univariate Heavy-Tailed Data*, ser. Series in Probability and Statistics. Wiley, 2007.

[49] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *arxiv :0706.1062*, Jun. 2007.

[50] B. R. Frieden, "Fisher information, disorder, and the equilibrium distributions of physics," *Physical Review A*, vol. 41, no. 8, pp. 4265–4276, 1990.

[51] B. R. Frieden, "Physics from Fisher Information : A Unification," *American Journal of Physics*, vol. 68, no. 11, p. 1064, 2000.

[52] B. R. Frieden, *Science from Fisher Information : A Unification*. Cambridge University Press, 2004.

[53] M. T. Martin, F. Pennini, and A. Plastino, "Fisher's information and the analysis of complex signals," *Physics Letters A*, vol. 256, no. 2-3, pp. 173–180, 1999.

[54] A. Nagy, "Fisher information in density functional theory," *The Journal of Chemical Physics*, vol. 119, no. 18, pp. 9401–9405, Nov. 2003.

[55] L. Telesca, V. Lapenna, and M. Lovallo, "Fisher information analysis of earthquake-related geoelectrical signals," *Natural Hazards and Earth System Sciences*, vol. 5, no. 4, pp. 561–564, 2005.

[56] K. D. Sen, J. Antolín, and J. C. Angulo, "Fisher-Shannon analysis of ionization processes and isoelectronic series," *Physical Review A*, vol. 76, no. 3, p. 32502, 2007.

[57] M. Balasco, V. Lapenna, M. Lovallo, G. Romano, A. Siniscalchi, and L. Telesca, "Fisher information measure analysis of earth's apparent resistivity," *International Journal of Nonlinear Science*, vol. 5, no. 3, pp. 230–236, 2008.

[58] S. A. Frank, "Natural selection maximizes Fisher information," *Journal of Evolutionary Biology*, vol. 22, no. 2, pp. 231–244, 2009.

[59] J. M. Borwein, A. S. Lewis, M. N. Limber, and D. Noll, "Maximum entropy reconstruction using derivative information part 2 : computational results," *Numerische Mathematik*, vol. 69, no. 3, pp. 243–256, 1995.

[60] J. M. Borwein, A. S. Lewis, and D. Noll, "Maximum entropy reconstruction using derivative information, part 1 : Fisher information and convex duality," *Mathematics of Operations Research*, vol. 21, pp. 442–468, 1996.

[61] V. Zivojnovic and D. Noll, "Minimum Fisher information spectral analysis," in *ICASSP-97, IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 1997, pp. 3957–3960.

[62] V. Zivojnovic, "Minimum Fisher information of moment-constrained distributions with application to robust blind identification," *Signal Processing*, vol. 65, no. 2, pp. 297–313, Mar. 1998.

[63] C. Vignat and J.-F. Bercher, "On Fisher information inequalities and score functions in non-invertible linear systems," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 4, p. Article 71, 2003. PDF

[64] C. Vignat and J. F. Bercher, "Analysis of signals in the Fisher-Shannon information plane," *Physics Letters A*, vol. 312, no. 1-2, pp. 27–33, Jun. 2003. PDF

[65] J. Bercher and C. Vignat, "On minimum Fisher information distributions with restricted support and fixed variance," *Information Sciences*, vol. 179, no. 22, pp. 3832–3842, Nov. 2009. PDF

[66] A. Dembo, T. Cover, and J. Thomas, "Information theoretic inequalities," *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1501–1518, 1991.

[67] A. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information and Control*, vol. 2, no. 2, pp. 101–112, 1959.

[68] N. Blachman, "The convolution inequality for entropy powers," *IEEE Transactions on Information Theory*, vol. 11, no. 2, pp. 267–271, 1965.

[69] V. Papathanasiou, "Some characteristic properties of the Fisher information matrix via Cacoullos-Type inequalities," *Journal of Multivariate Analysis*, vol. 44, no. 2, pp. 256–265, Feb. 1993.

[70] R. Zamir and M. Feder, "A generalization of the entropy power inequality with applications," *IEEE Transactions on Information Theory*, vol. 39, pp. 1723—1728, 1993.

[71] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *IEEE Transactions on Information Theory*, vol. 44, pp. 1246—1250, 1998.

[72] R. Zamir, "A necessary and sufficient condition for equality in the matrix Fisher-Information-Inequality," Tech. Rep. EE-S-97-11, 1997.

[73] C. Vignat and J.-F. Bercher, "Matrix Fisher inequalities for non-invertible linear systems," in *Proceedings of the IEEE International Symposium on Information Theory, 2002*, 2002, p. 237.

[74] E. Romera and J. S. Dehesa, "The Fisher-Shannon information plane, an electron correlation tool," *The Journal of Chemical Physics*, vol. 120, no. 19, pp. 8906–12, May 2004.

[75] J. Szabó, K. Sen, and A. Nagy, "The Fisher-Shannon information plane for atoms," *Physics Letters A*, vol. 372, no. 14, pp. 2428–2430, Mar. 2008.

[76] J. Sanudo and R. Lopez-Ruiz, "Alternative evaluation of statistical indicators in atoms : the non-relativistic and relativistic cases," *0904.1144*, Apr. 2009.

[77] H. Montgomery and K. Sen, "Statistical complexity and Fisher-Shannon information measure of h+2," *Physics Letters A*, vol. 372, no. 13, pp. 2271–2273, Mar. 2008.

[78] J. Antolín and J. C. Angulo, "Complexity analysis of ionization processes and isoelectronic series," *International Journal of Quantum Chemistry*, vol. 109, no. 3, pp. 586–593, 2009.

[79] J. Dehesa, P. Sánchez-Moreno, and R. Y. nez, "Cramer-Rao information plane of orthogonal hypergeometric polynomials," *Journal of Computational and Applied Mathematics*, vol. 186, no. 2, pp. 523–541, Feb. 2006.

[80] E. Romera and A. Nagy, "Fisher-Rényi entropy product and information plane," *Physics Letters A*, vol. 372, no. 46, pp. 6823–6825, Nov. 2008.

[81] G. Ferri, F. Pennini, and A. Plastino, "LMC-complexity and various chaotic regimes," *Physics Letters A*, vol. 373, no. 26, pp. 2210–2214, Jun. 2009.

[82] B. R. Frieden, *Physics from Fisher Information*. Cambridge University Press, 1998.

[83] R. B. Frieden and R. A. Gatenby, *Exploratory Data Analysis Using Fisher Information*, 1st ed. Springer, Nov. 2006.

[84] E. Uhrmann Klingen, "Minimal Fisher information distributions with compact-supports," *Sankhya : The Indian Journal of Statistics*, vol. 57, no. 3, pp. 360–374, 1995.

[85] E. Uhrmann-Klingen, "Fisher-Minimale Dichten auf Kompacten intervallen (in german)," Ph.D. dissertation, Essen, 1992.

[86] J.-F. Bercher and C. Vignat, "On minimum Fisher information distributions with restricted support and fixed variance," *Accepted in Information Sciences*, july 2009. PDF

[87] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. Cambridge University Press, 1927.

[88] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, Aug. 1991.

[89] P. Hammad, "Mesure d'ordre α de l'information au sens de Fisher," *Revue de Statistique Appliquée*, vol. 26, no. 1, pp. 73–84, 1978.

[90] J. Rufin, *Globalia*. Editions Gallimard, Jun. 2005.

[91] K. Iniewski, *Wireless Technologies : Circuits, Systems, and Devices*, 1st ed. CRC, Oct. 2007.

[92] L. Kahn, "Single-Sideband transmission by envelope elimination and restoration," *Proceedings of the IRE*, vol. 40, no. 7, pp. 803–806, 1952.

[93] C. Berland, J.-F. Bercher, I. Hibon, M. Villegas, D. Belot, D. Pache, V. Le Goascoz, "US Patent US2006034391A1 : Digital Transmitter Architecture," Feb. 2006. PDF

[94] R. B. Staszewski and P. T. Balsara, *All-Digital Frequency Synthesizer in Deep-Submicron CMOS*. Wiley-Interscience, Sep. 2006.

[95] R. Staszewski, J. Wallberg, S. Rezeq, C. Hung, O. Eliezer, S. Vemulapalli, C. Fernando, K. Maggio, R. Staszewski, N. Barton, M. Lee, P. Cruise, M. Entezari, K. Muhammad, and D. Leipold, "All-digital PLL and transmitter for mobile phones," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2469–2482, 2005.

[96] C. Joubert, J.-F. Bercher, G. Baudoin, T. Divel, S. Ramet, and P. Level, "Time behavioral model for phase-domain ADPLL based frequency synthesizer," in *Radio and Wireless Symposium, 2006 IEEE*, 2006, pp. 167–170. PDF

[97] C. Joubert, J.-F. Bercher, and G. Baudoin, "Contributions to the analysis and design of an ADPLL," in *13th IEEE International Conference on Electronics, Circuits and Systems, 2006. (ICECS) 2006.*, 2006, pp. 322–325. PDF

[98] G. Baudoin, J.-F. Bercher, and C. Joubert, "Calcul théorique du bruit de phase dû au dithering dans un oscillateur contrôlé numériquement," in *Journées Nationales Microondes 2007*, may 2007. PDF

[99] R. Staszewski, J. Wallberg, C. Hung, G. Feygin, M. Entezari, and D. Leipold, "LMS-based calibration of an RF digitally controlled oscillator for mobile phones," *IEEE Transactions on Circuits and Systems II : Express Briefs*, vol. 53, no. 3, pp. 225–229, 2006.

[100] P. Guillot, C. Berland, J.-F. Bercher, and P. Philippe, "A 2GHz 65nm CMOS digitally-tuned BAW oscillator," in *International Conference on Electronics Circuits and Systems (ICECS2008)*, Malta, Sep. 2008. PDF

[101] P. Guillot, P. Philippe, C. Berland, J. Bercher, and P. Gamand, "Low-noise high-resolution BAW-based high-frequency oscillator," *Electronics Letters*, vol. 45, no. 17, pp. 914–916, 2009. PDF

[102] P. Guillot, P. Philippe, C. Berland, and J.-F. Bercher, "Faisabilité de référence haute fréquence pour les architectures RF," in *Journées Nationales Microondes 2009*, may 2009. PDF

[103] J. Groe, "Polar transmitters for wireless communications," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 58–63, 2007.

[104] J. Groe, "A multimode cellular radio," *IEEE Transactions on Circuits and Systems II : Express Briefs*, vol. 55, no. 3, pp. 269–273, 2008.

[105] M. Markovic and H. Modi, "Feasibility of EER transmitters for 3GPP applications," in *IEEE Annual Wireless and Microwave Technology Conference, WAMICON '06*, 2006, pp. 1–2.

[106] W.-F. Loke, M.-W. Chia, and P.-Y. Chee, "Design considerations for multi-band OFDM polar transmitter of UWB system," *Electronics Letters*, vol. 43, pp. 3142–3145, 2007.

[107] B. Priyanto, T. Sorensen, O. Jensen, T. Larsen, T. Kolding, and P. Mogensen, "Impact of polar transmitter imperfections on UTRA LTE uplink performance," in *Norchip, 2007*, 2007, pp. 1–4.

[108] K. Waheed, R. B. Staszewski, and S. Rezeq, "Curse of digital polar transmission : Precise delay alignment in amplitude and phase modulation paths," in *IEEE International Symposium on Circuits and Systems, 2008. ISCAS 2008*, 2008, pp. 3142–3145.

[109] J.-F. Bercher and C. Berland, "Envelope and phase delays correction in an EER radio architecture," *Analog Integrated Circuits and Signal Processing*, vol. 55, pp. 21–35, Apr. 2008. PDF

[110] J. Bercher, A. Diet, C. Berland, G. Baudoin, and M. Villegas, "Monte-Carlo estimation of time mismatch effect in an OFDM EER architecture," in *Radio and Wireless Conference, 2004 IEEE*, 2004, pp. 283–286.

[111] J.-F. Bercher and C. Berland, "Envelope/phase delays correction in an EER radio architecture," in *13th IEEE International Conference on Electronics, Circuits and Systems, 2006. ICECS '06.*, 2006, pp. 443–446. PDF

[112] J.-F. Bercher and C. Berland, "Adaptive delays alignment in polar transmitter architecture," *submitted*, 2009. PDF

[113] J.-F. Bercher and C. Berland, "Adaptive time mismatches identification and correction in polar transmitter architecture," in *Wireless Technologies, 2007 European Conference on*, 2007, pp. 78–81.

[114] C. Berland, J.-F. Bercher, and O. Venard, "Digital signal processing techniques to compensate for RF imperfections in advanced transmitter architectures," in *European Wireless Technology Conference (EuWit2008)*, Amsterdam, Oct. 2008. PDF

[115] P. Hammad, *Information, systèmes et distribution*. Editions Cujas, Jul. 1987.

Chapitre 8

Recueil de publications

8.1 J.-F. Bercher, "Maximum entropy with fluctuating constraints - The example of K-distributions," *Physics Letters A*, vol. 372, no. 24, pp. 4361–4363, Jun. 2008.

Physics Letters A 372 (2008) 4361–4363

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Maximum entropy with fluctuating constraints: The example of K-distributions

J.-F. Bercher [a,b,*,1]

[a] *Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette cedex, France*
[b] *Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne la Vallée cedex 2, France*

ARTICLE INFO

Article history:
Received 17 January 2008
Received in revised form 20 March 2008
Accepted 3 April 2008
Available online 11 April 2008
Communicated by A.R. Bishop

PACS:
05.40.-a
02.50.-r
05.90.+m

Keywords:
Maximum entropy principle
Superstatistics
K-distributions

ABSTRACT

We indicate that in a maximum entropy setting, the thermodynamic β and the observation constraint are linked, so that fluctuations of the latter imposes fluctuations of the former. This gives an alternate viewpoint to 'superstatistics'. While a Gamma model for fluctuations of the β parameter gives the so-called Tsallis distributions, we work out the case of a Gamma model for fluctuations of the observable, and show that this leads to K-distributions. We draw attention to the fact that these heavy-tailed distributions have high interest in physical applications, and we discuss them in some details.

1. Introduction

The technique of compounding, or mixing, is a standard technique for the interpretation or construction of distributions [1, Chapter 8]. A random variable may result of the mixing of different populations, or the source of the variable may be unknown and the possible sources characterized by a probability density function. This technique has been employed in [2,3] for recovering the Tsallis (Generalized Pareto) distribution and generalized as 'Superstatistics'. This concept that has met interest [4–9] consists in varying the natural parameter β of a Boltzmann distribution according to some density $f_\beta(\beta)$ for β. In particular, the Tsallis distribution is associated with a Gamma distribution. Interestingly, this distribution had already been derived in such a way in [10].

The objective of this Letter is twofold. First, we relate fluctuations of the intensive parameter β to fluctuations of an observable $\bar{E}$, and indicate that the relevant distribution of the underlying system can be viewed as the maximum entropy solution with randomized constraints. Second, we work out the case of a Gamma distribution for the observable and draw attention to the physically relevant solution, a K-distribution, that is characterized and discussed.

2. Maximum entropy with fluctuating constraints

Consider a Gibbs ensemble where each system of the ensemble is composed of equilibrated subsystems, cells, each of them being described by the standard canonical factor $e^{-\beta E}$. The corresponding distribution can be regarded as the distribution that maximizes the Shannon–Boltzmann entropy

$$S = -\int f_E(x) \log f_E(x)\, d\Gamma \tag{1}$$

subject to normalization and to the observation constraint $\bar{E} = \int x f_E(x)\, d\Gamma$, and where $d\Gamma$ is the infinitesimal volume in the phase space. Resolution leads to $f_E(x|\bar{E}) = e^{-\beta x - \log Z(\beta)}$, where $Z(\beta) = \int e^{-\beta E}\, d\Gamma$ is the partition function and where β, the Lagrange multiplier associated to the observation constraint is solution of

$$S(\bar{E}) = \max_\beta \{-\beta \bar{E} - \log Z(\beta)\}. \tag{2}$$

This relation simply indicates that the entropy $S(\bar{E})$ and the potential $\Phi(\beta) = \log Z(\beta)$ are conjugated functionals while $\bar{E}$ and β are conjugated variables. This also gives the standard relationship

* Correspondence address: Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette cedex, France.
E-mail address: jf.bercher@esiee.fr.
[1] On sabbatical leave from ESIEE-Paris, France.

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2008.04.016

$$\bar{E} = -\frac{d\Phi(\beta)}{d\beta} = -\frac{d\log Z(\beta)}{d\beta}. \tag{3}$$

Hence, it is stressed that these variables are not independent, and that variations of the thermodynamic β necessarily implies variations of its dual variable $\bar{E}$ and reciprocally. Thus, if the mean energy varies from subsystem to subsystem (from cell to cell), and if these variations are modeled using a statistical distribution $f_{\bar{E}}(\bar{E})$, then the global distribution of the system under consideration will be

$$P(E) = \int_0^{+\infty} f_E(E|\bar{E}) f_{\bar{E}}(\bar{E})\, d\bar{E}, \tag{4}$$

where $f_E(x|\bar{E})$ is the Boltzmann distribution with mean $\bar{E}$. Given a model $f_{\bar{E}}(\bar{E})$ of fluctuations of $\bar{E}$ and the relationship (3) linking β and $\bar{E}$, we readily obtain the distribution for β:

$$f_\beta(\beta) = f_{\bar{E}}\left(\frac{-d\Phi(\beta)}{d\beta}\right)\left|\frac{d^2\Phi(\beta)}{d\beta^2}\right|. \tag{5}$$

Then, the system distribution can also be written

$$P(E) = \int_0^{+\infty} f_E(E|\beta) f_\beta(\beta)\, d\beta. \tag{6}$$

Of course this is similar to the 'Superstatistics' of Beck and Cohen [3], but with the supplementary ingredient of the relationship between the intensive parameter β and its dual variable $\bar{E}$, so that a model of fluctuations can be naturally introduced for $\bar{E}$. At this point, it can also be noted that the dimension of β is the inverse of an energy E.

3. Illustration in the canonical case

Let us now illustrate this point of view with a very simple example: the probability distributions of velocity and kinetic energy for a macroscopic body consisting of a large number of atoms. In the Gibbs distribution, the energy $E(p,q)$ can be decomposed in the sum of the kinetic and potential energy $E(p,q) = K(p) + U(q)$, so that $e^{-\beta E(p,q)}\, dp\, dq = e^{-\beta K(p)} e^{-\beta U(q)}\, dp\, dq$, with p the momenta and q the coordinates. Then the probability for the momenta reduces to $Ae^{-\beta K(p)}\, dp$. We know that for an atom of mass m the kinetic energy is $K(p) = (p_x^2 + p_y^2 + p_z^2)/2m$, where p_x, p_y, p_z are the coordinates of the momentum, and one obtains the classical Maxwellian distribution. Expressing the distribution in terms of the velocities, with $p = mv$ and $v^2 = v_x^2 + v_y^2 + v_z^2$, we have

$$f_V(v|\beta) = 4\pi\left(\frac{\beta m}{2\pi}\right)^{3/2} v^2 e^{-\beta m v^2/2} \tag{7}$$

for the magnitude of the velocity, and

$$f_E(E|\beta) = \frac{2}{\sqrt{\pi}} \beta^{\frac{3}{2}} \sqrt{E} e^{-\beta E} \tag{8}$$

for the magnitude of the kinetic energy, which is a χ^2 distribution with three degrees of freedom. If the motion was restricted to a single axis, we would have a single degree of freedom and an exponent $1/2$ for β instead of $3/2$ in the previous relations. The partition function is given by

$$Z(\beta) = \int_0^{+\infty} e^{-\beta K(p)}\, dp = \left[\frac{\pi}{2\beta m}\right]^{\frac{3}{2}}, \tag{9}$$

and we immediately obtain

$$\bar{E} = -\frac{d}{d\beta}\log Z(\beta) = \frac{3}{2\beta}. \tag{10}$$

Therefore, in the canonical case, the mean value of the kinetic energy is simply proportional to the inverse of the parameter β. Fluctuations of this mean value $\bar{E}$ can be modeled through a distribution $f_{\bar{E}}$.

4. The example of K-distributions

It is well known, e.g. [2,10], that Gamma fluctuations of the 'inverse temperature' β leads to a Tsallis distribution. Because of the relationship (10), this corresponds to an inverse-Gamma distribution for $\bar{E}$. Another reasonable model for fluctuations of the mean energy $\bar{E}$ is a Gamma distribution since it proves highly flexible for modeling the distribution of positive variables. It also appears as the distribution of the sum of squared Gaussian variables.

4.1. Derivation

Therefore, according to (10), β is distributed according to the inverse-Gamma distribution: $f_\beta(\beta) = \frac{b^a}{\Gamma(a)}\beta^{-a-1}e^{-\frac{b}{\beta}}$. Then, $P(E)$ is given by

$$P(E) = \frac{2}{\sqrt{\pi}}\frac{b^a}{\Gamma(a)} \int_0^{+\infty} \beta^{\frac{3}{2}} \sqrt{E} e^{-\beta E} \beta^{-a-1} e^{-\frac{b}{\beta}}\, d\beta. \tag{11}$$

Formula [3.471.9] in [11] enables to perform the integration with respect to β. This formula reads $\int_0^{+\infty} x^{\nu-1} e^{-\frac{\beta}{x}-\gamma x} dx = 2(\frac{\beta}{\gamma})^{\frac{\nu}{2}} K_\nu(2\sqrt{\beta\gamma})$ with $\beta, \gamma > 0$, and where K_ν is a modified Bessel function the of second kind of order ν. This leads to

$$P(E) = \frac{2}{E}\frac{1}{\Gamma(a)\Gamma(3/2)}(bE)^{\frac{3/2+a}{2}} K_{-\frac{3}{2}+a}(2\sqrt{bE}). \tag{12}$$

As far as the velocity is concerned, the variable change $E = \frac{1}{2}mv^2$ gives

$$P_V(v) = \frac{4}{\Gamma(3/2)\Gamma(a)} v^{a+\frac{1}{2}} c^{a+\frac{3}{2}} K_{a-\frac{3}{2}}(2cv) \tag{13}$$

with $c = \sqrt{bm/2}$.

4.2. Main properties and characterizations

These distributions are known as K-distributions. They are extensively used for modeling the statistics of interferences of radiowaves, radar clutter, optical scintillation, ultrasound scattering, etc. [12–19]. Interestingly, a recent application of inverse-chi-square superstatistics in medical statistics has been proposed [20] with excellent agreement with real data. The shape of distribution (12) and its evolution when the parameter a varies is given in Fig. 1. This distribution is derived in radar imaging as the mixture of a Rayleigh and Gamma distribution [18] and competes with log-normal e.g. [15] or generalized Pareto (Tsallis) distributions e.g. [17] for fitting observed data. It is worth mentioning that the estimation of the parameters of this distribution is well documented, e.g. [14,16,18].

The case of general mixture of Gamma distributions was considered in [21] where is given a formula similar to (12). A K-distribution was also obtained in [4] in the limit case of a ratio of two Gamma distributions (that is a F-distribution) for describing fluctuations of β. A model of inverse chi-square for β was mentioned and its asymptotics predicted in [22]. These asymptotics are easy to observe: formula [9.7.2] in [23] gives the limit form for the Bessel K function for large arguments: $K_\nu(z) \approx \sqrt{\frac{\pi}{2z}} e^{-z}$. Consequently, $P(E)$ behaves as

$$P(E) \sim \frac{2}{\Gamma(a)}(bE)^{\frac{a-1}{2}} e^{-2\sqrt{bE}}, \tag{14}$$

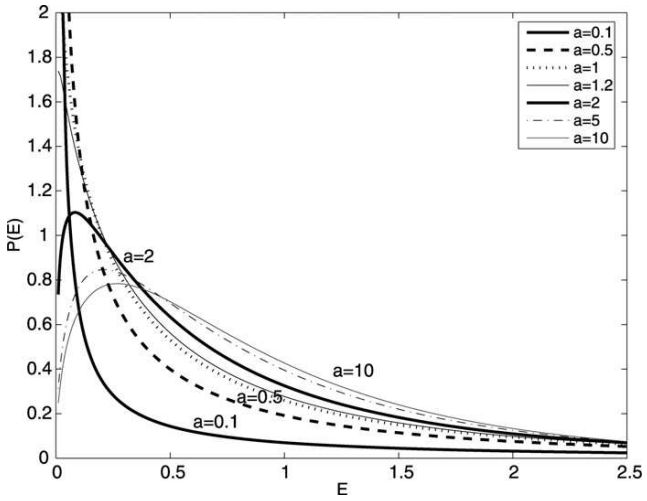


Fig. 1. Examples of K-distributions (12) with $a = 0.1$ to $a = 10$, and with $b = 3a/2$ chosen so that the expectation of all distributions is equal to 1.

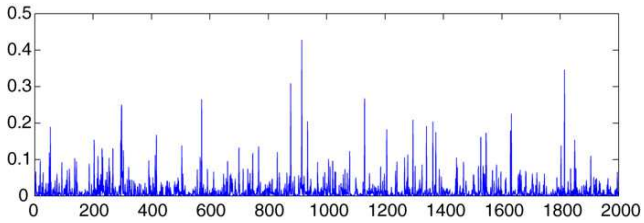


Fig. 2. A typical realization of a K-distribution with $a = 0.5$.

and presents an exponential tail. The survival function $\bar{F}(x) = \Pr(X \geqslant x)$ then has the limit form $\bar{F}(x) = \frac{1}{\Gamma(a)b}(\frac{1}{2})^{a-1}\Gamma(a+1, \alpha 2\sqrt{bx})$, where $\Gamma(a+1, 2\sqrt{bx})$ is the upper incomplete Gamma function. It behaves as $(2\sqrt{bx})^a e^{-2\sqrt{bx}}$ for large values of x and $\lim_{x\to+\infty} e^{\lambda x}\bar{F}(x) = +\infty$ for all positive λ, which means that the distribution is heavy-tailed.

For small values of the argument, [23] again gives a limiting form, formula [9.6.9]: $K_\nu(z) \approx \frac{1}{2}\Gamma(\nu)(\frac{1}{2}z)^{-\nu}$ for $\nu > 0$. For $a > 3/2$ we then have

$$P(E) \sim \frac{\Gamma(a-\frac{3}{2})}{\Gamma(a)\Gamma(3/2)} b^{\frac{3}{2}}\sqrt{E}. \tag{15}$$

The case $a < 3/2$ follows using the fact that $K_{-\nu}(z) = K_\nu(z)$:

$$P(E) \sim \frac{\Gamma(\frac{3}{2}-a)}{\Gamma(a)\Gamma(3/2)} b^a E^{a-1}. \tag{16}$$

This last expression shows that the K-distribution presents an integrable singularity at the origin for $a < 1$, and explains the behavior shown in Fig. 1. In such a case, there is a preferential weighting of small values interrupted by some rare large events, an interesting feature encountered in many experiments. A typical realization obtained with $a = 0.5$ is reported in Fig. 2.

As a final comment, let us recall that the distribution of the ratio of a Gaussian variable and a Gamma variable is a Tsallis distribution. Along the same line, a stochastic model of K-distributed variates can be a simple multiplicative model. Indeed, if z is the product of two Gamma variables x and y, the distribution of z is $p_Z(z) = \int p_X(x)p_Y(z/x)\,dx/x$, which reduces to a K-distribution.

5. Conclusion

This Letter presents two contributions. First, we have highlighted the fact that within a maximum entropy setting, the β parameter and observable are conjugated, and linked by relation (3). Therefore, their fluctuations are also linked. This gives an alternate viewpoint to 'Superstatistics', since the model of fluctuations can be introduced for the observable rather than for the intensive parameter.

As an illustration and second contribution, we worked out the natural example of a Gamma distribution, and exhibited the system distribution in the form of a K-distribution. We underlined that this interesting distribution is highly relevant in some physical applications, and have emphasized some of its properties.

References

[1] N. Johnson, S. Kotz, A. Kemp, Univariate Discrete Distributions, second ed., John Wiley, New York, 1992.
[2] G. Wilk, Z. Wlodarczyk, Phys. Rev. Lett. 84 (2000) 2770.
[3] C. Beck, E. Cohen, Physica A 322 (2003) 267.
[4] F. Sattin, L. Salasnich, Phys. Rev. E 65 (2002) 035106.
[5] A.G. Bashkirov, Physica A 340 (2004) 153.
[6] S.M.D. Queirós, Europhys. Lett. 71 (3) (2005) 339.
[7] T. Yamano, Prog. Theor. Phys. Suppl. 162 (2006) 87.
[8] G. Wilk, Z. Wlodarczyk, Physica A: Stat. Theor. Phys. 376 (2007) 279.
[9] S. Abe, C. Beck, E.G.D. Cohen, Phys. Rev. E 76 (2007) 031102.
[10] B.A. Maguire, E.S. Pearson, H.A. Wynn, Biometrika 39 (1952) 168.
[11] I. Gradshteyn, I. Ryzhik, Table of Integrals, Series and Products, Academic Press, 1965.
[12] E. Jakeman, P.N. Pusey, Phys. Rev. Lett. 40 (9) (1978) 546.
[13] E. Jakeman, J. Phys. A 13 (1980) 31.
[14] R.S. Raghavan, IEEE Trans. Aerospace Electronic Systems 27 (2) (1991) 238.
[15] A. Abdi, M. Kaveh, Electron. Lett. 34 (1998) 851.
[16] D.R. Iskander, A.M. Zoubir, B. Boashash, IEEE Trans. Signal Proc. 47 (1999) 1147.
[17] B. LaCour, IEEE J. Oceanic Eng. 29 (2004) 310.
[18] K.D. Ward, R.J.A. Tough, S. Watts, Sea Clutter: Scattering, the K Distribution and Radar Performance, Institution for Engineering and Technology, 2006.
[19] L. Weng, J.M. Reid, P.M. Shankar, K. Soetanto, J. Acoust. Soc. Am. 89 (1991) 2992.
[20] L.L. Chen, C. Beck, Physica A 387 (2008) 3162.
[21] S.K. Bhattacharya, Metrika 11 (1966) 133.
[22] H. Touchette, C. Beck, Phys. Rev. E 71 (2005) 016131.
[23] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables, Dover Publications, 1965.

8.2 J.-F. Bercher and C. Vignat, "Estimating the entropy of a signal with applications," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1687–1694, 2000.

IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 48, NO. 6, JUNE 2000 1687

Estimating the Entropy of a Signal with Applications

Jean-François Bercher and Christophe Vignat

***Abstract*—In this paper, we present a new estimator of the entropy of continuous signals. We model the unknown probability density of data in the form of an AR spectrum density and use regularized long-AR models to identify the AR parameters. We then derive both an analytical expression and a practical procedure for estimating the entropy from sample data. We indicate how to incorporate recursive and adaptive features in the procedure. We evaluate and compare the new estimator with other estimators based on histograms, kernel density models, and order statistics. Finally, we give several examples of applications. An adaptive version of our entropy estimator is applied to detection of law changes, blind deconvolution, and source separation.**

***Index Terms*—AR processes, entropy estimation, parametric methods, regularization, spectrum analysis.**

I. Introduction

SINCE Shannon's work [1], entropy is used as a major tool in information theory. However, this tool is rarely used in signal processing, except in theoretical frameworks, because it appears difficult to compute or estimate the entropy from a set of real data. Interesting approaches involving direct use of entropy for signal processing applications can be found in [2]–[4]. In many applications, a measure of complexity of underlying probability density functions, or a measure of dependence between components or signals, allows the design of an optimal processing scheme, possibly in nonstationary contexts. Examples of such situations are plentiful:

- source separation;
- blind deconvolution;
- source coding;
- image alignment;
- detection of abrupt changes;

and so on. Thus, entropy-based approaches might be useful for such problems.

The entropy $H(X)$ of a random variable X with continuous probability density function (PDF) $p_X(x)$ is defined as

$$H(X) = -E_X[\log_2 p_X] = -\int_{-\infty}^{+\infty} p_X(x)\log_2 p_X(x)\,dx. \tag{1}$$

In the discrete case, where X takes values x_i with probabilities p_i, $H(X) = -\sum_i p_i \log_2 p_i$. Basic estimates can be built, using any raw estimates of the p_i in the preceding formula. More sophisticated entropy estimates, based on coding theorems that are specific to the discrete case, can be found in [5] and [6]. In signal processing, data usually have continuous PDF because of contamination by continuous noise. In this continuous case, two main approaches exist. First, the PDF can be approximated by an element of a parameterized set, whose entropy is known in term of the parameters [7]. Second, entropy estimators are based on a prior estimation of underlying PDF's (or cumulative distribution functions) using methods such as histograms [4], [8], order statistics [9], [10] (see [11] for a comparative study), or kernel methods [2], [3], [8].

In this paper, we derive and apply to signal processing problems a new estimator of entropy for all continuous PDF's with bounded support. This estimator can be implemented in recursive schemes and has tracking capabilities in nonstationary contexts. Furthermore, our approach provides a convenient estimation procedure of PDF's.

The main contributions of this work are

i) a new presentation and improvements of the approach of [12]–[14] for PDF estimation;
ii) derivation of an analytical close-form formula for the estimation of entropy;
iii) presentation of a practical procedure for entropy estimation, including recursive and adaptive features;
iv) evaluation and comparisons with other methods;
v) examples of application of the entropy estimator to signal processing problems.

This paper is organized as follows. In Section II, we discuss the relevance of AR modeling of PDF's and introduce regularized long-AR models. In Section III, we give the theoretical expression of the entropy associated with AR-PDF's and a practical procedure for estimating the entropy. In Section IV, we give examples of PDF's estimation. Then, we analyze and compare the behavior of the new estimator with other methods. Finally, in Section V, we give some applications of this estimate to signal processing problems, namely, detection of PDF changes, blind equalization, and source separation.

II. AR Modeling of PDF's

A. Introduction

Our approach consists of estimating the unknown PDF $p_X(x)$ as the power density spectrum $S_W(x)$ of some unit variance AR process $W(\omega, n)$. Applications of spectral estimation methods to PDF estimation were first introduced in the context of nonlinear signal processing in [12], [15]. AR-PDF estimation was also discussed in [13] and [14]. We discuss here the relevance of this model for PDF estimation, recall the link between spectral matching and linear prediction, and then propose to use a

Manuscript received May 17, 1999; revised December 10, 1999. The associate editor coordinating the review of this paper and approving it for publication was Prof. Jian Li.

J.-F. Bercher is with the Laboratoire Signaux et Télécoms, Groupe ESIEE, Noisy-le-Grand, France (e-mail: bercherj@esiee.fr).

C. Vignat is with the Laboratoire Systèmes de Communications, Université de Marne la Vallée, Noisy-le-Grand, France (e-mail: vignat@univ-mlv.fr).

Publisher Item Identifier S 1053-587X(00)04069-1.

long-AR regularized approach in order to obtain stable and accurate estimates. In the sequel, we will show that this model leads to an easy procedure for computing the associated entropy and that recursivity and adaptivity can be introduced in the procedure.

We suppose that the observation consists in samples of a process $X(\omega,n)$, identically distributed according to a continuous PDF $p_X(x)$ with bounded support, say, $[-(1/2),+(1/2)]$. This hypothesis, although restrictive, is usual in the context of PDF estimation.

B. PDF Estimation Using AR Modeling

We look for an estimate $\hat{p}_X(x)$ of the true (unknown) PDF $p_X(x)$ parameterized by a set of coefficients $\{a_k\}_{1\le k\le p}$ in the form of a power spectrum density $S_W(x)$

$$\hat{p}_X(x)=S_W(x)=\frac{\sigma_\epsilon^2}{\left|1-\sum_{k=1}^{p}a_k e^{-j2\pi kx}\right|^2} \tag{2}$$

where σ_ϵ^2 is chosen such that $\int_{-(1/2)}^{+(1/2)} S_W(f)\,df=1$.

The relevance of this parameterization lies in the fact that any continuous spectrum density can be approximated, in the $\|\cdot\|_\infty$ sense, by an AR spectrum density. More precisely, if $S_Z(x)$ is a symmetric continuous spectral density on $[-(1/2),(1/2)]$, and $\delta>0$, then there exists an integer p and a real-valued causal AR(p) process $W(\omega,n)$ with innovation variance σ_ϵ^2 such that $\|S_W(x)-S_Z(x)\|_\infty<\delta$ (see [16, corol. 4.4.2. p. 132]). This result extends easily to the case of nonsymmetric, possibly one-sided, $S_Z(x)$; in this case, $W(\omega,n)$ is a complex-valued AR process.

Once the analogy between PDF's and power spectrum densities is stated, a natural question arises: Can we find a process $Z(\omega,n)$ whose spectrum is precisely the PDF of the random variable $X(\omega)$? It is easy to check that $Z(\omega,n)=e^{j(nX+\phi(\omega))}$ has this property if X is any sample of process $X(\omega,n)$, and $\phi(\omega)$ is uniformly distributed over $[0,2\pi]$ and independent of X. Indeed, its correlation function $R_Z(k)$ is nothing but the first characteristic function of X.

However, this "underlying process" $Z(\omega,n)$ of $X(\omega,n)$ is very likely not an AR process. Hence, to identify the parameters $\{a_k\}$ associated with PDF $\hat{p}_X$ in (2), we need to match a given spectrum $S_Z(x)=p_X(x)$ with an AR spectrum $S_W(x)=\hat{p}_X(x)$. A classical result about spectral matching [17] states that the best AR(p) model spectrum minimizing the integrated ratio of the two spectra $I(Z,W)=\int_{-(1/2)}^{+(1/2)} S_Z(x)/S_W(x)\,dx$ is nothing but the AR solution of the linear prediction problem whose parameters $\mathbf{a}=[a_1,\ldots,a_p]^T$ are such that $\mathbf{R}_Z\mathbf{a}=\mathbf{r}_Z$.[1] Matrix $(\mathbf{R}_Z)_{1\le i,j\le p}=R_Z(i-j)$ and correlation vector $(\mathbf{r}_Z)_{1\le i\le p}=R_Z(i)$ are built using correlation function $R_Z(k)$.

Thus, modelization of PDF p_X as an AR spectrum follows the two following steps: 1) estimation of the correlation sequence $R_Z(k)$, i.e., of the characteristic function of X, using the available data $\{x(n)\}_{1\le n\le N}$ from $X(\omega,n)$, as the statistical average correlation estimate

$$\hat{R}_Z(k)=\frac{1}{N}\sum_{n=1}^{N}e^{jkx(n)}$$

and 2) estimation of the coefficients $\{a_k\}$ of the AR process $W(\omega,n)$ by solving $\hat{\mathbf{R}}_Z\mathbf{a}=\hat{\mathbf{r}}_Z$.

C. Long AR Models and Regularization

As mentioned above, real PDF's are very likely not in the form of an AR spectrum. Hence, an accurate modelization of PDF's via AR techniques may require the use of long AR models. However, the counterpart of adopting a high number of coefficients is a loss in the stability of the estimate (e.g., spurious peaks). The exploitation of regularization techniques enables the use of long AR models, and thus, modeling of "non-AR" spectra, while preserving stability.

The idea is to use a long AR model with the addition of some prior knowledge about the "smoothness" of the spectrum. In [20], Kitagawa and Gersch defined the kth smoothness by

$$D_k=\int_0^1\left|\frac{\partial^k A(f)}{\partial f^k}\right|^2 df$$

with $A(f)=\sum_{k=1}^{p}a_k e^{j2\pi kf}$, and showed that $D_k\propto\mathbf{a}^t\mathbf{\Delta}_k\mathbf{a}$, where $\mathbf{\Delta}_k$ is the diagonal matrix with elements $[\mathbf{\Delta}_k]_{ii}=i^{2k}$.

The AR parameters are obtained as a regularized least-squares solution

$$\hat{\mathbf{a}}=(\hat{\mathbf{R}}_Z+\lambda\mathbf{\Delta}_k)^{-1}\hat{\mathbf{r}}_Z \tag{3}$$

where hyperparameter λ balances a fidelity to the data and a smoothness prior.

In [20] and [21], a Bayesian interpretation of this regularized least-squares is derived, which also leads to a selection rule for the hyperparameter λ, as the minimizer of the following marginal likelihood:

$$L(\lambda)=\log(\det(\hat{\mathbf{R}}_Z+\lambda\mathbf{\Delta}_k))-p\log(\lambda)-N\log(\sigma_Z^2) \tag{4}$$

where σ_Z^2 is chosen such that the AR probability distribution is properly normalized.

Let us now turn to the problem of computing an estimate $\hat{H}(X)$ of the entropy $H(X)$ associated with $p_X(x)$. A natural approach at this step is to build the entropy estimate $\hat{H}(X)$ of the unknown PDF $p_X(x)$ as the entropy of the estimate PDF $\hat{p}_X(x)$.

III. Estimate of Entropy

In this section, we exhibit the analytical expression of the entropy associated with $\hat{p}_X(x)$. Then, we give an alternate and easier procedure for estimating the entropy. Finally, we show how to introduce recursivity and adaptivity in the procedure.

A. Theoretical Expression

The exact expression of entropy $\hat{H}(X)$, using $\hat{p}_X(x)$ defined as in (2) can be derived. Let us denote by $\{z_k\}$ the set of p supposedly simple poles of $\hat{p}_X(x)$ and by $\{\mu_k\}$ the set of associated

[1]Note that minimizing $\log I(Z,W)$ is equivalent to the minimization of $I(Z,W)$ and, in the case of a "good matching" [17], to the maximization of the Burg entropy. Hence, the general AR spectral matching method coincides with the "maximum entropy spectral estimation method," which was derived in the case of gaussian signals [18], [19].

residues. With $A(z) = \sum_{k=1}^{p} a_k z^{-k}$, a straightforward but tedious calculus (omitted here) yields (5), shown at the bottom of the page. Direct use of this analytical expression of entropy is obviously difficult since it requires computation of all poles of the AR model, together with their respective residues, which can be highly time consuming, particularly in the case of long AR models. Although we have here an explicit formula (5), it is desirable to find an equivalent formula for $\hat{H}(X)$ that provides an easier estimation procedure.

B. Easier Estimation Procedure

The entropy associated with $\hat{p}_X(x)$ as defined by (2) is $\hat{H}(X) = -\int_{-(1/2)}^{+(1/2)} S_W(x) \log_2 S_W(x)\, dx$. Hence, applying the Plancherel–Parseval formula to the right-hand side of the above relation yields

$$\begin{aligned}\hat{H} &= -\sum_{k=-\infty}^{+\infty} R_W(k) C_W^*(k) \\ &= -2\,\mathrm{Re}\left\{\sum_{k=0}^{+\infty} R_W(k) C_W^*(k)\right\} \end{aligned} \tag{6}$$

where $R_W(k)$ denotes the kth correlation coefficient of $W(\omega, n)$, and $C_W(k) = FT^{-1}[\log_2 S_W(x)]$ denotes the kth component of its cepstrum. Note that both $R_W(k)$ and $C_W(k)$ have Hermitian symmetry since $S_W(x)$ is real, which provides the right-hand side of (6).

At this step, we take advantage of the AR structure of process $W(\omega, n)$ since, for that particular type of process, both correlation and cepstrum functions obey recursive relations [22][2]:

$$R_W(k) = \sum_{i=1}^{p} a_i R_W(k-i) + \sigma_\epsilon^2 \delta(k) \tag{7}$$

$$C_W(k) = \begin{cases} \log \sigma_\epsilon^2, & \text{if } k = 0 \\ h(k) - \displaystyle\sum_{i=1}^{k-1} \left(\frac{i}{k}\right) C_W(i) h(k-i), & \text{if } k > 0 \end{cases} \tag{8}$$

with $h(k)$ the impulse response of the AR system, which is also computed recursively according to

$$h(k) = -\sum_{i=1}^{p} a_i h(k-i) + \delta(k). \tag{9}$$

The estimated entropy can thus be computed using (6)–(9), avoiding any numerical integration. Obviously, however, in practice, the infinite sum in (6) should be truncated, which leads to some truncation error. It is also important to note that (6) does not require the explicit estimation of the PDF but only of the first $(p+1)$ coefficients of the correlation sequence $R_W(k)$, which in turn enables the computation of the AR parameters and the cepstrum involved in (6).

As will be described below, as part of a practical method for implementing the method expressed by (6)–(9), it is possible to estimate recursively the correlation sequence.

C. Implementation in a Recursive Scheme

The first step consists of estimating the characteristic function R_W from the observation data $\{x_i\}_{1 \le i \le n+1}$. The statistical average correlation sequence can be estimated recursively using

$$\begin{aligned} R_W^{(n+1)}(k) &= \frac{1}{n+1} \sum_{i=1}^{n+1} e^{j2\pi k x_i} \\ &= \frac{n}{n+1} R_W^{(n)}(k) + \frac{1}{n+1} e^{j2\pi k x_{n+1}}. \end{aligned} \tag{10}$$

This empirical characteristic function is the inverse Fourier transform of the empirical distribution $\hat{p}_X^{(n+1)}(x) = 1/(n+1)\sum_{i=1}^{n+1} \delta(x - x_i)$. In kernel methods for density estimation, the empirical distribution is smoothed using a kernel $\phi(x)$. It is also possible to compute such an estimate recursively, as in (10); see [23] and references therein. For the characteristic function, with $\Phi(k) = \mathrm{FT}^{-1}(\phi(x))$, this leads to

$$R_W^{(n+1)}(k) = \frac{n}{n+1} R_W^{(n)}(k) + \frac{1}{n+1} \Phi(k) e^{j2\pi k x_{n+1}}. \tag{11}$$

The proposed method thus consists of the three following steps:

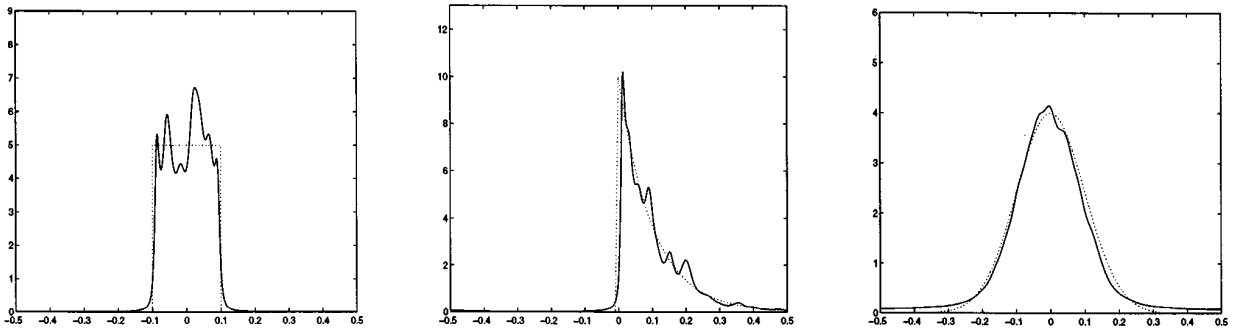
- *First step* Estimate the $(p+1)$ correlation coefficients $R_W^{(n+1)}(k)_{0 \le k \le p}$ using the $(n+1)$ available samples using (10) or (11).
- *Second step* The set of estimated correlations $R_W^{(n+1)}(k)_{0 \le k \le p}$ allows computation of parameters $a_{i\ 1 \le i \le p}^{(n+1)}$ and, thus, time series $R_W^{(n+1)}(k)$ and $C_W^{(n+1)}(k)$, using relations (7)–(9).
- *Third step* Finally, application of (6) gives the estimated entropy of process $X(\omega, n)$ based on its $(n+1)$ first samples.

Furthermore, it is also straightforward to derive an adaptive version of this entropy estimation scheme. It suffices to introduce a forgetting factor μ in the updating formula (10) of the correlation sequence. For the correlation matrix, this gives

$$\hat{\mathbf{R}}_W^{(n)} = \frac{1}{n}\left[(n-1)\mu \hat{\mathbf{R}}_W^{(n-1)} + \mathbf{e}(n)\mathbf{e}(n)^+\right]$$

[2]Relation (8) is derived using [22] and the fact that $R_W(k) = \sigma_\epsilon^2 h(k) * h(-k)^*$, where $h(k)$ is minimum phase with $h(0) = 1$.

$$\hat{H}(X) = \frac{I_2}{I_1} + \log I_1, \qquad \text{with} \begin{cases} I_1 = \sigma_\epsilon^{-2} = \displaystyle\sum_{k=1}^{p} \mu_k^* \left(1 - A\left(\frac{1}{z_k^*}\right)\right)^{-1} \\ I_2 = 2\,\mathrm{Re}\left\{\displaystyle\sum_{k=1}^{p} \mu_k^* \left(1 - A\left(\frac{1}{z_k^*}\right)\right)^{-1} \log\left\{1 - A\left(\frac{1}{z_k^*}\right)\right\}\right\}. \end{cases} \tag{5}$$



(a) (b) (c)

Fig. 1. AR estimates and (a) theoretical uniform (b) exponential, and (c) Gaussian laws.

with $\mathbf{e}(n)^+ = [e^{-j(x_n+\phi)} \ \dots \ e^{-j(px_n+\phi)}]$. Then, the AR parameters and entropy can be evaluated at each new sample using (3) and (6)–(9).

It is also possible to compute recursively the AR parameters, thus avoiding the matrix inversion required in (3). Indeed, the regularized least squares solution can be computed recursively, using a gradient approach [24]

$$\mathbf{a}^{(n+1)} = \mathbf{a}^{(n)} + \alpha \left[\left(\hat{\mathbf{R}}_W^{(n)} + \lambda \boldsymbol{\Delta}_k \right) \mathbf{a}^{(n)} - \hat{\mathbf{r}}^{(n)} \right]. \quad (12)$$

Finally, it is also possible to adopt a more simple "LMS-like" approach, such as

$$\mathbf{a}^{(n+1)} = \mathbf{a}^{(n)} + \alpha \left[(\mathbf{e}(n)\mathbf{e}(n)^+ + \lambda \boldsymbol{\Delta}_k)\mathbf{a}^{(n)} - e^{-j\phi}\mathbf{e}(n) \right]. \quad (13)$$

IV. Simulation Results and Comparisons

A. AR-PDF Estimation

In order to illustrate the versatility of the long AR approach for PDF estimation, experiments were performed on sequences of 250 samples distributed according to

a) a uniform PDF $U_{[-(1/10),(1/10)]}$;
b) an exponential PDF with parameter 0.1;
c) a Gaussian PDF $\mathcal{N}(0, 0.01)$.[3]

For these three PDF's, the parameters p and λ are, respectively,

a) $p = 32$, $\lambda = 5\ 10^{-5}$;
b) $p = 32$, $\lambda = 3\ 10^{-5}$;
c) $p = 20$, $\lambda = 8\ 10^{-4}$.

Results given in Fig. 1(a)–(c) show the relevance of this approach, which is able to approximate with accuracy different shapes of PDF's.

B. Entropy Estimation

In order to analyze the behavior of the AR entropy estimator, we performed a Monte Carlo study in the case of a uniform and a Gaussian PDF. We evaluated the mean and standard deviation σ_H of the AR entropy estimate over 50 realizations as a function of the length of available data. We compared these results with those obtained in the cases of histogram and kernel PDF approximation as well as the modified Vasicek's estimator recommended in [11].

- In the case of histograms, the entropy is estimated as $-\sum_{i=1}^{K} N_i/K \log_2(N_i/K)$, where N_i is the number of values in the ith bin, and K is the number of bins.
- In the kernel approach, the available samples $\{x_1, \dots x_N\}$ are directly used for modeling the density as $\hat{p}_X(x) = (1/N)\sum_{i=1}^{N} \phi(x - x_i)$, where $\phi(x)$ is a smoothing kernel, which is usually chosen as a Gaussian kernel. In our experiments, we evaluated the density on a grid of $L = 1000$ points with a kernel width chosen to provide the best results. Finally, the entropy was evaluated as $-\sum_{i=1}^{L} \hat{p}_i \log_2 \hat{p}_i$.
- The Vasicek's estimator [9] relies on the remark that $H(X) = \int_0^1 \log_2(dF^{-1}(u)/du)\, du$. Then, the estimator is obtained by approaching the cumulative distribution function F with order statistics. The modified Vasicek's estimator [11] has the following form:

$$V_{m,n}(X) = \frac{1}{n}\sum_{i=1}^{n} \log_2 \left(\frac{n}{2m}\left(x_{(i+m)} - x_{(i-m)}\right) \right) + f(m,n)$$

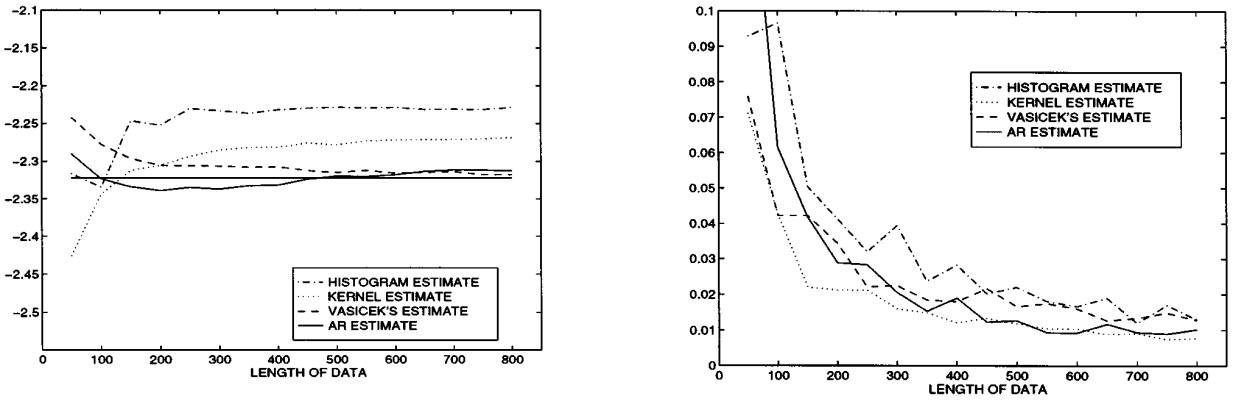
where

$\{x_{(i)}\}$ ordered set of samples $x(i)$ (with $x_{(i)} = x_{(1)}$ for $i < 1$ and $x_{(i)} = x_{(n)}$ for $i > n$);
m positive integer;
$f(m,n)$ function that accounts for a bias correction; see [11].

Figs. 2 and 3 give the results [mean (a) and standard deviation (b)] for a uniform density $U_{[-(1/10),(1/10)]}$ and for a normal density $\mathcal{N}(0, 0.01)$, with respectively theoretical entropy $H(X) = -2.3219$ bits and $H(X) = -1.2748$ bits. For the uniform density, we have chosen an AR order $p = 32$, 64 bins for histogram estimates, $\sigma^2 = 5$ (on the grid of 1000 points) for kernel estimates, and $m = 3$ for the modified Vasicek's estimator. Results for the Gaussian density were obtained with $p = 20$, 20 bins for histogram estimates, $\sigma^2 = 10$ for kernel estimates, and $m = 3$ for the modified Vasicek's estimator.

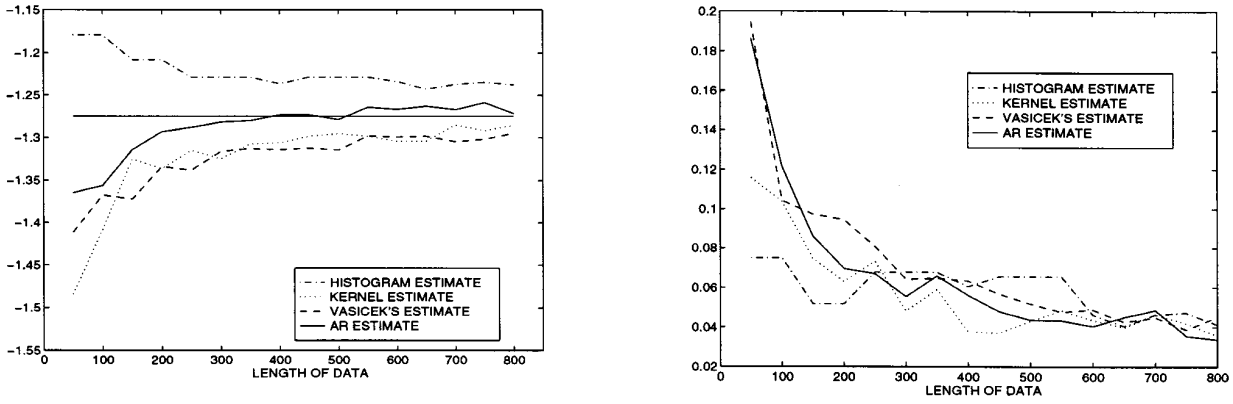
These results exhibit the good statistical behavior of our estimator, that is, a low bias and a small variance. For the uniform density, the AR-based estimator has about the same performance as the Vasicek's one concerning the bias but a lower standard deviation. The other estimators present a higher bias. In the case of the Gaussian density, the AR-based estimator clearly

[3]Since PDF's are modeled as power spectra on interval $[-(1/2), +(1/2)]$, the data had to be scaled on this interval. This does not restrict our approach because the entropy of the scaled variable differs from the original entropy only by a known additive term.



(a) (b)

Fig. 2. Comparison of histogram, kernel, Vasicek's, and AR-based estimates of entropy for a uniform density. (a) Mean values. (b) Standard deviations.



(a) (b)

Fig. 3. Comparison of histogram, kernel, Vasicek's, and AR-based estimates of entropy for a Gaussian density. (a) Mean values. (b) Standard deviations.

presents the lowest bias, whereas all have about the same standard deviation. Considering these two test cases, the AR-based estimator of entropy shows either comparable or lower bias and variance. Hence, it proves accurate and compares favorably with all other estimators considered.

As far as the complexity is concerned, the AR-based estimator also compares favorably with others since it suffices to estimate some correlation coefficients, find parameters using a gradient recursion, construct two times series, and compute their scalar product. The kernel approach requires a large amount of memory in order to evaluate and store the density estimate, even in the case of recursive kernel estimators (storage requirements can be reduced by using a coarser grid adjusted to the data range, but the bias increases with the length of intervals on the grid). Entropy evaluation requires L multiplications and L evaluations of $\log_2$. Vasicek's estimator requires storing and sorting the data and n evaluations of $\log_2$.

V. Sample Applications

A. Detecting PDF Changes

An interesting application of the adaptive estimates of Section III-C consists of detecting PDF changes in signals. As an illustration, we consider a signal $x(n)$ that is composed of 200 samples generated according to a mixture of two Gaussian distributions, with means ± 0.3 and standard deviation $\sigma = 0.06$, followed by 200 samples distributed uniformly on the interval $[-0.44, 0.44]$ and by 200 samples of the same gaussian mixture. First PDF has entropy $H_1 = -1.011$ bits, whereas the second has entropy $H_2 = -0.1844$ bits.

Fig. 4 shows signal $x(n)$; it is difficult, by a simple inspection, to diagnose that there are PDF changes. Fig. 5 shows the adaptive estimates [computed using (12) and (13)] of the negentropy of this test signal, using a forgetting factor $\mu = 0.98$.

The following points are of importance.

i) PDF changes appear clearly.
ii) Rupture points are properly revealed.
iii) The entropy is estimated with accuracy.
iv) The adaptive estimate has a good tracking capability.

B. Blind Deconvolution of AR Systems

The problem of blind deconvolution consists of recovering the input X and possibly the parameters of a filter from the sole observation of its output Y. The concept of entropy brings an interesting answer to this problem [25], relying on the following proposition.

Proposition 1: Let $Y(\omega, n)$ be the output of a unit norm filter whose input is a non-Gaussian i.i.d. sequence $X(\omega, n)$. Then, $H(Y) > H(X)$.

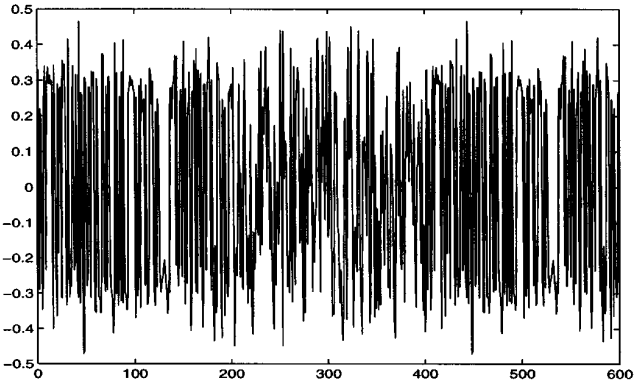


Fig. 4. Test signal for adaptive estimates: 200 samples from a gaussian mixture, 200 samples from a uniform distribution and 200 samples from a gaussian mixture.

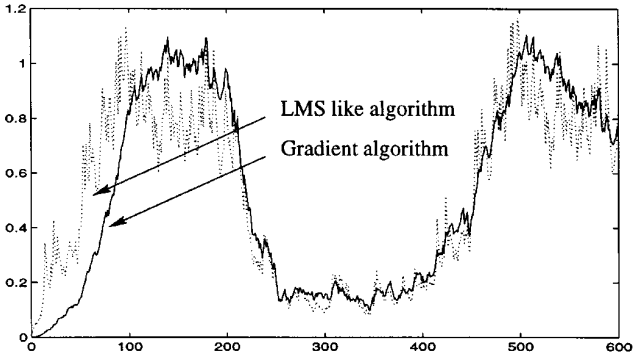


Fig. 5. Adaptive estimates of negentropy for signal in Fig. 4.

The intuitive reason behind this result is that $p_Y(y)$ is closer to a Gaussian distribution than $p_X(x)$, where the Gaussian distribution has the maximum entropy in the set of distributions of given variance; see [25].

Proof: The key of the proof is the entropy power inequality stated by Shannon [1, Th. 15 and App. 6]; see also [26].

Entropy power inequality: *If X and Y are two independent random variables with entropies $H(X)$ and $H(Y)$, then*

$$e^{2H(X+Y)} \geq e^{2H(X)} + e^{2H(Y)}$$

with equality if and only if X and Y are Gaussian variables.

The term "entropy power" comes from the fact that the power of a Gaussian variable X is proportional to $e^{2H(X)}$.

Let g be the impulse response of the filter with input $X(\omega, n)$. Its output $Y(\omega, n)$ is $Y(\omega, n) = \sum_i g_i X(\omega, n-i)$. The classical result on the entropy of rescaled variables and the assumption of stationarity give

$$\begin{aligned} H(g_i X(\omega, n-i)) &= H(X(\omega, n-i)) + \log_2 |g_i| \\ &= H(X) + \log_2 |g_i|. \end{aligned}$$

Now, the entropy power inequality gives

$$\begin{aligned} e^{2H(Y)} &\geq \sum_i e^{2H(g_i X(\omega, n-i))} = \sum_i e^{2H(X)+2\log_2 |g_i|} \\ &= e^{2H(X)} \frac{1}{\log_e 2} \sum_i |g_i|^2. \end{aligned}$$

Then, taking $\log_e$ of both sides, we have

$$H(Y) \geq H(X) + \frac{1}{2} \log_2 \sum_i |g_i|^2.$$

Finally, for a unit norm filter, that is, $\sum_i |g_i|^2 = 1$, the last relation reduces to $H(Y) \geq H(X)$. □

The deconvolution procedure then simply consists of adjusting the parameters θ of a filter F_θ, with input Y, such that its output $\hat{X}$ has minimum (estimated) entropy. If θ are the filter parameters, this becomes

$$\theta_{\text{opt}} = \arg\min_\theta \hat{H}(\hat{X})$$
$$\text{submitted to } \begin{cases} \hat{X}(f) = F_\theta(f) Y(f) \\ ||F_\theta||^2 = 1. \end{cases}$$

Simulations were performed in the case of nonminimum–phase AR filters. They showed that the AR parameters can be identified very accurately and that the input can be perfectly reconstructed, even if the AR order is overestimated. These simulations were performed in the case of uniform and binary inputs, with 500 samples of data, and the initial solution was chosen as a standard minimum-phase solution.

In the case of non-AR filters, experiments showed that the procedure suffers from local minima. However, the procedure may prove of value when used in a compound criterion that should be considered in the presence of observation noise, such as

$$[\theta_{\text{opt}}, \mathbf{X}_{\text{opt}}] = \arg\min_{\theta, \hat{\mathbf{X}}} \hat{H}(\hat{X}) + \alpha ||\mathbf{Y} - \mathbf{H}_\theta \hat{\mathbf{X}}||^2$$

where $\mathbf{H}_\theta$ is a convolution matrix. Note also that the previous criterion can be used in a standard deconvolution context, where $\mathbf{H}_\theta$ is known. In this case, the entropy term will help to select the "right" solution among several equivalent solutions, as it occurs in ill-posed problems.

C. Source Separation

In the context of source separation, N signals $\mathbf{s}(n) = [s_1(n), \ldots, s_N(n)]$ are mixed by an unknown $N \times N$ matrix $\mathbf{A}$ to provide observed signals $\mathbf{x}(n) = [x_1(n), \ldots, x_N(n)]$. The problem consists of recovering the sources from the sole observation of signals $\mathbf{x}(n)$ using only an assumption on the mutual independence of sources. The objective is reached by designing a matrix $\mathbf{B}$ such that the reconstructed signal $\hat{\mathbf{s}}(n) = \mathbf{B}\mathbf{x}(n)$ has independent components. The information theoretic measure of independence is the mutual information I_M, that is, the Kullback–Leibler divergence between $p_{\hat{S}_1, \ldots, \hat{S}_N}(\hat{s}_1, \ldots, \hat{s}_N)$ and $\Pi_{1 \leq i \leq N} p_{\hat{S}_i}(\hat{s}_i)$. For the source separation problem, minimization of the mutual information reduces to the minimization of

$$C(\mathbf{B}) = -\log |\det \mathbf{B}| + \sum_{i=1}^{N} H(\hat{s}_i). \tag{14}$$

In classical approaches, as no estimate of entropy is available, $\mathbf{B}$ is chosen as the solution of the nonlinear decorrelation equations: $E[\hat{s}_i \psi_j(\hat{s}_j)] = 0$, for $i \neq j$, which express the stationarity condition of $C(\mathbf{B})$. Function $\psi_j(s_j)$, which is the so-called score function, is the log derivative of the density $p_{S_j}(s_j)$.

TABLE I
EXPERIMENTS ON SOURCE SEPARATION. RESULTING MATRIX $M = AB$ SHOULD BE THE IDENTITY.

source 1	source 2	$M = AB$
$U_{[-0.5,0.5]}$	$\frac{1}{2}\{N(0.3,0.2) + N(-0.3,0.2)\}$	$\begin{bmatrix} 1 & 0.0395 \\ 0.0011 & 1 \end{bmatrix}$
$U_{[-0.5,0.5]}$	$U_{[-0.5,0.5]}$	$\begin{bmatrix} 1 & 0.02 \\ 0.0163 & 1 \end{bmatrix}$
binary $[-\frac{1}{2},\frac{1}{2}]$	$N(0,1)$	$\begin{bmatrix} 1 & 0.0354 \\ 0.02 & 1 \end{bmatrix}$
$U_{[-0.5,0.5]}$	$N(0,1)$	$\begin{bmatrix} 1 & 0.23 \\ -0.0491 & 1 \end{bmatrix}$

Using our AR parameterization, we can i) either estimate the cost function $C(\mathbf{B})$ and minimize it using any standard optimization procedure or ii) estimate the solution of the decorrelation equations using the analytical expression (in terms of the AR parameters) of the score function $\psi(x) = \mathbf{a}^+\mathbf{T}\mathbf{a}/|\mathbf{a}^+\mathbf{e}|^2$, where $\mathbf{T}$ is the Toeplitz matrix with entries $\mathbf{T}_{kl} = -j2\pi(k-l)e^{-j2\pi(k-l)x}$, and $\mathbf{e}^+ = [1\, e^{j2\pi x} \ldots e^{j2\pi(p-1)x}]$.

We performed simulations using the first approach, using 500 samples of data in the case of the mixture of $N = 2$ sources. Table I presents, for several distributions of the sources, the resulting matrix $M = AB$, which should be the identity matrix, up to a scaling factor and a permutation. These results show that this approach enables proper separation of the input sources, with performances comparable with classical methods [3].

VI. CONCLUSION

The concept of entropy plays a central role in information theory. However, it is rarely used directly in signal processing applications. In this paper, we have presented an estimator of the entropy of a signal and illustrated its behavior through several motivating examples of signal processing applications.

Our estimator relies on a simple analogy between the problems of PDF estimation and power spectrum estimation. The problem of PDF estimation is tackled using an AR modelization, which is a well-known approach in signal processing. This parametric modelization enables the accurate description of a large class of PDF. Moreover, in order to obtain accurate and stable estimates, we have chosen to use the long AR approach of [20], where the problem of AR parameters estimation is regularized by a smoothness constraint. Relying on the AR modelization, we have presented an estimation procedure for the entropy in a recursive scheme. The corresponding estimator does not require the explicit estimation of the PDF but only of some samples of a correlation sequence. Thus, it is easy to derive adaptive versions of this estimator. As illustrated by a simulation study, the AR estimate of entropy proves accurate and compares favorably with other classical estimates.

Finally, we have given several examples of applications where the entropy-based approach provides valuable results. It is worth recalling that many other applications can be considered, as soon as they involve some measure of dependence between random variables or signals or a measure of complexity.

ACKNOWLEDGMENT

The authors wish to thank A. Pagès-Zamora for useful comments on an early version of this manuscript and sending us [12] and a preprint of [15]. Suggestions and comments of F. Barbaresco were helpful and are acknowledged. The authors also wish to thank the anonymous referees for their valuable comments and suggestions.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423; 623–656, July/Oct. 1948 [Online] Available http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html.
[2] P. Viola, N. N. Schraudolph, and T. J. Sejnowski, "Empirical entropy manipulation for real-world problems," in *Advances in Neural Information Processing Systems 8*. Cambridge, MA: MIT Press, 1996.
[3] D. T. Pham, "Blind separation of instantaneous mixture of sources via an independent component analysis," *IEEE Trans. Signal Processing*, vol. 44, pp. 2768–2779, Nov. 1996.
[4] R. Moddemeijer, "On estimation of entropy and mutual information of continuous distributions," *Signal Process.*, vol. 16, no. 3, pp. 233–246, 1989.
[5] I. Kontoyiannis *et al.*, "Nonparametric entropy estimation for stationary processes and random fields, with application to english text," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1319–1327, May 1998.
[6] D. S. Ornstein and B. Weiss, "Entropy and data compression schemes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 78–83, Jan. 1993.
[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[8] P. Hall and S. Morton, "On the estimation of entropy," *Ann. Inst. Stat. Math.*, vol. 45, pp. 69–88, 1993.
[9] O. Vasicek, "A test of normality based on sample entropy," *J. R. Stat. Soc. Ser. B*, vol. 38, pp. 54–59, 1976.
[10] J. C. Correa, "A new estimator of entropy," *Commun. Stat.–Theory Methodol.*, vol. 24, pp. 2439–2449, 1995.
[11] R. Wieczorkowski and P. Grzegorzewski, "Entropy estimators—Improvements and comparisons," *Commun. Stat.–Simul. Comput.*, vol. 28, no. 2, pp. 541–567, 1999.
[12] A. Pagès-Zamora and M. A. Lagunas, "New approaches in nonlinear signal processing: Estimation of the PDF function by spectral estimation methods," in *Proc. IEEE-Athos Workshop Higher-Order Stat.*, June 1995, pp. 204–208.
[13] S. Kay, "Model-based probability density function estimation," *IEEE Signal Processing Lett.*, vol. 5, pp. 318–320, Dec. 1998.
[14] J.-F. Bercher and C. Vignat, "Estimating the entropy of a signal with applications," in *Proc. ICASSP*, Phoenix, AZ, Mar. 1999.
[15] A. Pagès-Zamora and M. A. Lagunas, "Fourier models for nonlinear signal processing," *Signal Process.*, vol. 76, no. 1, pp. 1–16, 1999.
[16] P. J. Brockwell and R. A. Davis, *Times Series: Theory and Methods*, 2nd ed. New York: Springer-Verlag, 1987.
[17] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.
[18] J. P. Burg, "Maximum entropy spectral analysis," presented at the Proc. 37th Meet. Soc. Explor. Geophys., 1967.
[19] J. E. Shore, "Minimum cross-entropy spectral analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 230–237, Apr. 1981.
[20] G. Kitagawa and W. Gersh, "A smoothness priors long AR model method for spectral estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 57–65, 1985.
[21] J.-F. Giovannelli, G. Demoment, and A. Herment, "A bayesian method for long AR spectral estimation: A comparative study," *IEEE Trans. Ultrason. Freq. Ferroelect.*, vol. 43, pp. 220–233, Mar. 1996.
[22] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[23] O. Hössjer and U. Host, "On-line density estimators with high efficiency," *IEEE Trans. Inform. Theory*, vol. 41, pp. 829–835, May 1995.

[24] S. Haykin, *Adptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[25] D. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis II*. New York: Academic, 1981, pp. 565–609.
[26] A. Dembo and T. M. Cover, "Information theoretic inequalities," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1501–1518, Nov. 1991.

Jean-François Bercher was born in Marseille, France, in 1965. He received the B.E.E. degree in 1990 from the Institut National Polytechnique de Grenoble, Grenoble, France. He then received the Ph.D. degree in physics from the Laboratoire des Signaux et Systèmes, Université de Paris-Sud, Orsay, France, in 1995.

He is now with the Laboratoire Signaux et Télécoms of the ESIEE, Noisy-le-Grand, France, where he is Associate Professor. His interests are in applications of information theory to inverse problems, signal processing, and telecommunications.

Dr. Bercher is affiliated with MGEN.

Christophe Vignat was born in France in 1965. He received the B.E.E. degree from the École Supérieure d'Électricité, Paris, France. He then received the Ph.D. degree in physics from the Laboratoire des Signaux et Systèmes, the Université de Paris-Sud, Orsay, France, in 1993.

Since 1995, he has been Associate Professor at the Université de Marne-la-Vallée, Noisy-le-Grand, France. After working on adaptive systems, his interests shifted toward applications of signal processing to communication systems.

Dr. Vignat is a member of CROUS.

8.3 J.-F. Bercher, "Tsallis distribution as a standard maximum entropy solution with 'tail' constraint," *Physics Letters A*, vol. 372, no. 35, pp. 5657–5659, Aug. 2008.

Physics Letters A 372 (2008) 5657–5659

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Tsallis distribution as a standard maximum entropy solution with 'tail' constraint

J.-F. Bercher [a,b,*,1]

[a] *Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne la Vallée cedex 2, France*
[b] *Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette cedex, France*

ARTICLE INFO

Article history:
Received 7 February 2008
Received in revised form 13 June 2008
Accepted 18 June 2008
Available online 11 July 2008
Communicated by C.R. Doering

PACS:
89.70.Cf
02.50.-r
05.90.+m

Keywords:
Maximum entropy principle
Rényi entropy
Nonextensive statistics

ABSTRACT

We show that Tsallis' distributions can be derived from the standard (Shannon) maximum entropy setting, by incorporating a constraint on the divergence between the distribution and another distribution imagined as its tail. In this setting, we find an underlying entropy which is the Rényi entropy. Furthermore, escort distributions and generalized means appear as a direct consequence of the construction. Finally, the "maximum entropy tail distribution" is identified as a Generalized Pareto Distribution.

1. Introduction

The maximizers of a special entropy, the Tsallis entropy [1], with suitable constraints, are often called Tsallis distributions. It is worth mentioning that the maximization of any monotonous transform of Tsallis entropy, with the same constraints, leads to the same maximizers. This is in particular the case of Rényi entropy. In applied fields, Tsallis distributions (q-distributions) have encountered a large success because of their remarkable agreement with experimental data, see for instance [2,3] and references therein. These distributions are of very high interest in many physical systems, since they can exhibit heavy-tails, and model power-law phenomena. Indeed, power-laws are especially interesting since they appear widely in physics, biology, economy, and many other fields [4]. Tsallis distributions are similar to Generalized Pareto Distributions, which also have an high interest in other fields, namely reliability theory [5], climatology [6], radar imaging [7] or actuarial sciences [8]. Hence, a remarkable feature of the maximum Tsallis entropy construction is its ability to exhibit heavy tailed distributions. Furthermore, the essence of Pickands' extreme values theorem [9] is that the distribution of excesses over a threshold converges, under wide conditions, to a q-exponential [10]. Following the idea that an interest of Tsallis distributions is in fact *a tale of tails*, in the words of [11], we suggest that this kind of distributions can also be obtained from the familiar (Shannon) maximum entropy setting, by the introduction of an appropriate constraint. In this setting, we show that Rényi entropy appears naturally, and also obtain a natural interpretation of nonextensive escort distributions, 'generalized means' and entropic index.

2. Maximum entropy with 'tail' constraint

Jaynes' maximum entropy principle [12,13] suggests that the least biased probability distribution that describes a partially-known system is the probability distribution with maximum entropy compatible with all the available prior information. The Kullback–Leibler information divergence $D(P\|Q)$ measures the divergence of a distribution P to another distribution Q.

2.1. Problem setting

We call 'tail distribution' the probability density function (pdf) of the excesses X_u of a variate X over a threshold u. It is not only the tail part of the distribution, but is shifted by u toward the origin, and, as a pdf, is normalized to 1. Let now P and Q be two probability density functions. As a guideline, we find useful to imagine P as the tail distribution of Q: in such a case, these distributions are closely related. Our idea is to account more generally for an existing relationship between two distributions. Therefore,

* Correspondence address: Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne la Vallée cedex 2, France.
E-mail address: jf.bercher@esiee.fr.
[1] On sabbatical leave from ESIEE-Paris, France.

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2008.06.088

we propose to parametrize the relation between a candidate distribution and its 'parent' by the value θ of the Kullback–Leibler divergence between them: $D(P\|Q)=\theta$. This define a set of possible distributions, and the solution is selected in this set, according to the maximum entropy principle, as the distribution P with maximum entropy. This writes as follows:

$$\begin{cases} \max_P H(P) = -\int P(x)\log P(x)\,dx \\ \quad \text{s.t. } D(P\|Q) = \int P(x)\log\frac{P(x)}{Q(x)}dx = \theta. \end{cases} \tag{1}$$

The definition of $D(P\|Q)$ requires that P is absolutely continuous with respect to Q. It is understood, as usual, that $0\log 0 = 0\log 0/a = 0\log 0/0 = 0$ and that $a\log(a/0) = +\infty$ if $a>0$.

An alternative formulation, which leads to the same solution, could be to look for the distribution with minimum divergence to Q, in the set of all distributions with a given entropy:

$$\begin{cases} \min_P D(P\|Q) \\ \quad \text{s.t. } H(P) = \theta'. \end{cases} \tag{2}$$

It is clear that the optimum distribution arising from this procedure will not be the 'exact' tail distribution, since the distribution of the excesses $X_u = X-u|X>u$, which reads $X_u = X-u$ if $X>u$ (conditioned variable), has the pdf $P_u(x) = Q(x+u)/\bar{Q}(u)$, for $x\geqslant 0$, and where $\bar{Q}(u) = \int_u^{+\infty} Q(x)\,dx$ is the so-called survival function. It is rather the maximum entropy variant, when the "tail distribution" P is constrained to be at given divergence θ to the parent distribution Q. In fact, we do not need to rely on tails in this construction, but we simply introduce a constraint on the Kullback–Leibler divergence to a 'parent' distribution in order to account for a general relationship between these distributions.

2.2. The maximum entropy solution and first consequences

The solution of (1) is easily derived using standard maximum entropy results, e.g. [14]. In this case, we obtain

$$P(x) = \frac{1}{Z(\lambda)} e^{\lambda \log \frac{P(x)}{Q(x)}} \tag{3}$$

with λ the Lagrange parameter associated to the constraint $D(P\|Q)=\theta$, and $Z(\lambda)$ the partition function. It can also be reduced to

$$P(x) = \frac{1}{Z(\lambda)^{\frac{1}{1-\lambda}}} Q(x)^{\frac{\lambda}{\lambda-1}} \tag{4}$$

with

$$\log Z(\lambda) = (1-\lambda)\log\int Q(x)^{\frac{\lambda}{\lambda-1}}\,dx. \tag{5}$$

As usual in the maximum entropy setting, the constraint and partition function are linked by

$$\frac{d\log Z(\lambda)}{d\lambda} = \frac{d}{d\lambda}\log\int e^{\lambda\log\frac{P(x)}{Q(x)}}\,dx = \theta. \tag{6}$$

Let us now denote

$$q = \frac{\lambda}{\lambda-1}. \tag{7}$$

Then, it clearly appears on the one hand that the 'maximum entropy tail' (4), which can be rewritten as

$$P(x) = \frac{Q(x)^q}{\int Q(x)^q\,dx}, \tag{8}$$

is the *escort distribution* of nonextensive thermostatistics. On the other hand, the log-partition function (5) becomes

$$\log Z(q) = \frac{1}{1-q}\log\int Q(x)^q\,dx \tag{9}$$

where the right-hand side has exactly the form of the *Rényi entropy* of distribution Q. It shall be mentioned that the Rényi entropy has no definite concavity for $q>1$ and is not Lesche stable [15,16]. Its use in statistical physics has been discussed, e.g. [17]. In our context, it appears as a by-product of our original problem (1) which involves the maximization of the standard Boltzmann–Shannon entropy.

For the optimum distribution, we can also observe that the maximum (Shannon) entropy reduces to

$$H(P) = -\int P(x)\log P(x)\,dx \tag{10}$$

$$= -\int P(x)\left(\lambda\log\frac{P(x)}{Q(x)} - \log Z(\lambda)\right)dx \tag{11}$$

$$= -\lambda\int P(x)\log\frac{P(x)}{Q(x)}dx + \log Z(\lambda) \tag{12}$$

$$= -\lambda\theta + \log Z(\lambda) \tag{13}$$

where the last relation is obtained using the definition of θ in (1).

Therefore, we obtain that the maximum entropy problem (1) has for optimum value the Rényi entropy (9) with index q, minus a linear function of the constraint.

2.3. Solution with an additional observation constraint

Suppose now that the original problem is completed by an additional constraint. Indeed, the definition of equilibrium distributions usually need to take into account observation constraints. For instance, one often has to account for an observable defined as a mean value under distribution P. This is very classical in maximum entropy approaches. The constraint writes

$$m = \mathrm{E}_P[X]. \tag{14}$$

Since P is the escort distribution (8) of Q, this mean constraint is also the *generalized mean* constraint of nonextensive thermostatistics:

$$m = \int xP(x)\,dx = \frac{1}{\int Q(x)^q\,dx}\int xQ(x)^q\,dx. \tag{15}$$

When P is thought as the 'tail', m is the well-known *mean residual lifetime* (in reliability theory), or *expected future lifetime* (in survival analysis).

Because of the additional constraint, determination of the 'maximum entropy tail' distribution P amounts to further maximize the entropy in (13) subject to that constraint. Using (9), we obtain

$$\begin{cases} \max_Q \frac{1}{1-q}\log\int Q(x)^q\,dx \\ \quad \text{s.t. } \mathrm{E}_P[X] = m, \end{cases} \tag{16}$$

which leads to

$$Q(x) \propto \big(1-\beta(1-q)(x-m)\big)^{\frac{1}{1-q}} \tag{17}$$

for the value of the parameter β such that $\mathrm{E}_P[X]=m$, with P given by (8). As far as the latter is concerned, its expression is simply

$$P(x) \propto \big(1-\beta(1-q)(x-m)\big)^{\frac{q}{1-q}}. \tag{18}$$

This relation can also be rearranged as

$$P(x) \propto \big(1+\beta(q-1)(x-m)\big)^{-\frac{1}{q-1}-1}, \tag{19}$$

which is exactly in the form of the Generalized Pareto Distribution [18], with shape factor $(q-1)$ and scale factor β. Using (7), the exponent in (19) reduces to $-\lambda$, so that the distribution asymptotically behaves as a power-law with exponent $-\lambda$.

Observe that with $\theta = 0$ in (1), we readily have $P = Q$ and $q = 1$, and P in (19) reduces to the classical Boltzmann–Gibbs canonical distribution.

As a final comment, let us note that the maximum entropy tail distribution is *stable* with respect to thresholding, in the sense that the result remains in the same family, with entropic index q but a different scale parameter. If u is the threshold, then the pdf of excesses over u, with P the parent distribution, is

$$P_u(x) \propto P(x+u) \propto \big(1+\beta(q-1)(x+u-m)\big)^{-\frac{1}{q-1}-1} \propto \big(1+\beta'(q-1)(x-m)\big)^{-\frac{1}{q-1}-1}, \tag{20}$$

where the last term, with $\beta' = \beta/(1+\beta u)$, is obtained by factoring $(1+\beta u)$. This highlights the particular status of the 'maximum entropy tail distribution' as a tail distribution.

3. Conclusion

In this Letter, we followed two ideas. First, that an important feature of Tsallis' distributions is their ability to model the tail of distributions, particularly of those with heavy tails. The second idea is that these distributions should appear in a standard maximum entropy setting. This led us to the introduction of a constraint on the divergence between two distributions, one of them being imagined as the tail distribution. This constraint accounts for an existing general relationship between two distributions.

We showed that within this construction, the escort distributions and generalized means of nonextensive statistics appear very naturally. We also obtained that the original maximum (Shannon) entropy reduces to a maximum Rényi entropy; or equivalently to the maximization of Tsallis entropy. As far as the entropic index q is concerned, it is simply associated to the value of the divergence between the distribution Q and its escort distribution P. Finally, the 'maximum entropy tail' distributions, in the sense adopted in this Letter, are found to be Generalized Pareto Distributions, which have proved very useful for modeling heavy-tailed distributions in many applied problems. We believe that this construction can be useful to workers in the field.

References

[1] C. Tsallis, J. Stat. Phys. 52 (1988) 479.
[2] C. Beck, Physica A 331 (2004) 173.
[3] C. Tsallis, Chaos Solitons Fractals 13 (2002) 371.
[4] M.E.J. Newman, Contemp. Phys. 46 (2005) 323.
[5] M. Asadi, I. Bayramoglu, IEEE Trans. Rel. 55 (2006) 314.
[6] M.A.J.V. Montfort, J.V. Witter, Hydrological Sci. J. 31 (1986) 151.
[7] B. LaCour, IEEE J. Ocean. Eng. 29 (2004) 310.
[8] A.C. Cebrian, M. Denuit, P. Lambert, N. Am. Actuar. J. 7 (2003) 18.
[9] J. Pickands, Ann. Stat. 3 (1975) 119.
[10] J.-F. Bercher, C. Vignat, A new look at q-exponential distributions via excess statistics, doi:10.1016/j.physa.2008.05.038, Physica A (2008), in press.
[11] E.W. Montroll, M.F. Shlesinger, J. Stat. Phys. 32 (1983) 209.
[12] E.T. Jaynes, Phys. Rev. 108 (1957) 171.
[13] E.T. Jaynes, Proc. IEEE 70 (1982) 939.
[14] I. Csiszár, Ann. Stat. 19 (1991) 2032.
[15] B. Lesche, J. Stat. Phys. 27 (1982) 419.
[16] S. Abe, Phys. Rev. E 66 (2002) 46134.
[17] P. Jizba, T. Arimitsu, Phys. Rev. E 69 (2004) 026128, pMID: 14995541.
[18] N.L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, second ed., Wiley-Interscience, 1994.

8.4 J.-F. Bercher and C. Vignat, "A new look at q-exponential distributions via excess statistics," *Physica A*, vol. 387, no. 22, pp. 5422–5432, Sep. 2008.

Physica A 387 (2008) 5422–5432

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A new look at q-exponential distributions via excess statistics

J.-F. Bercher [a,b,*,1], C. Vignat [c]

[a] Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette Cedex, France
[b] Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France
[c] Université de Marne-la-Vallée, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France

ARTICLE INFO

Article history:
Received 25 February 2008
Received in revised form 6 May 2008
Available online 29 May 2008

PACS:
02.50.-r
05.40.-a
05.90.+m

Keywords:
q-exponentials
Generalized Pareto Distributions
Excess distributions
Maximum entropy principle
Nonextensive statistics

ABSTRACT

Q-exponential distributions play an important role in nonextensive statistics. They appear as the canonical distributions, i.e. the maximum generalized q-entropy distributions under mean constraint. Their relevance is also independently justified by their appearance in the theory of superstatistics introduced by Beck and Cohen. In this paper, we provide a third and independent rationale for these distributions. We indicate that q-exponentials are stable by a statistical normalization operation, and that Pickands' extreme values theorem plays the role of a CLT-like theorem in this context. This suggests that q-exponentials can arise in many contexts if the system at hand or the measurement device introduces some threshold. Moreover we give an asymptotic connection between excess distributions and maximum q-entropy. We also highlight the role of Generalized Pareto Distributions in many applications and present several methods for the practical estimation of q-exponential parameters.

1. Introduction

Since the pioneering paper [1] by Tsallis (1988), nonextensive statistics have received an increasing interest in the statistical physics community. As a consequence, the canonical distributions that appear in this context have led to many studies: these distributions are the q-Gaussians under the assumption that the total energy is fixed, and the q-exponentials when the mean value of the system is fixed.

In several cases and a wide variety of fields, experiments, numerical results and analytical derivations fairly agree with the formalism and the description by q-Gaussians or q-exponentials. This includes situations characterized by long-range interactions, long-range memory or space-time multifractal structure. Applications include fully developed turbulence, Levy anomalous diffusion, statistics of cosmic rays, econometry, and many others. Furthermore, recent studies have highlighted their excellent ability to approximate with high accuracy the exact distribution of complex systems [2].

Let us recall that q-exponential distributions are defined by

$$f_{q,\beta}(x) = \frac{1}{Z_{q,\beta}} e_q(-\beta x) \tag{1}$$

with

$$e_q(-\beta x) = (1 - \beta(1-q)x)^{\frac{1}{1-q}}, \quad x \in [0, A_q[\tag{2}$$

* Corresponding address: Laboratoire des Signaux et Systemes, CNRS-Univ Paris Sud-Supelec, 2, boulevard Blaise Pascal Citi DESCARTES BP 99, 91192 Gif-sur-Yvette Cedex, France.
E-mail addresses: jf.bercher@esiee.fr (J.-F. Bercher), vignat@univ-mlv.fr (C. Vignat).
[1] On sabbatical leave from ESIEE-Paris, France.

0378-4371/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2008.05.038

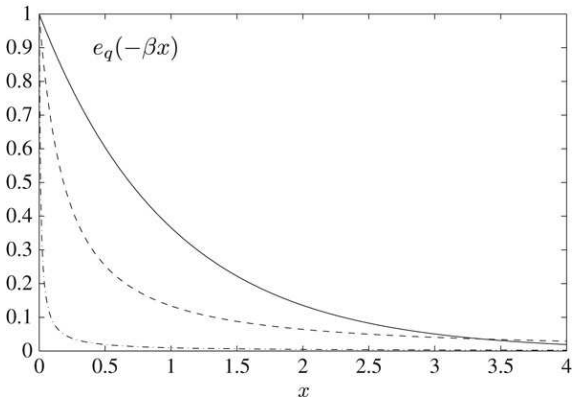


Fig. 1. Infinite support q-exponential distributions with $\beta = \frac{1}{2-q}$ and $q = 1.01$ (solid), 1.99 (dash-dot) and 1.8 (dash).

where A_q is the right endpoint, $A_q = \frac{1}{\beta(1-q)}$ if $q < 1$ and $A_q = +\infty$ if $1 < q < 2$, and where β is a positive scale parameter related to the mean m of $f_{q,\beta}$ as $m = \int x f_{q,\beta}(x)\,\mathrm{d}x = \frac{1}{\beta(3-2q)}$, for $q < 3/2$. The partition function $Z_{q,\beta}$ is

$$Z_{q,\beta} = \frac{1}{\beta\,(2-q)}, \quad 0 < q < 2.$$

We will also denote $\bar{F}_{q,\beta}(x)$ the so-called survival function associated to the q-exponential distribution (1):

$$\bar{F}_{q,\beta}(x) = \int_x^{+\infty} f_{q,\beta}(z)\mathrm{d}z \;=\; 1 - F_{q,\beta}(x) \tag{3}$$

$$= (1-\beta(1-q)x)^{\frac{1}{1-q}+1} \tag{4}$$

where $F_{q,\beta}(x)$ is the probability distribution function.

The q-exponential distributions $f_{q,\beta}(x)$ are the solutions of the following maximum entropy problem:

$$\max_f H_q(f) \quad \text{with} \int_0^{+\infty} x f^q(x)\,\mathrm{d}x = \theta \quad \text{and} \quad \int_0^{+\infty} f(x)\,\mathrm{d}x = 1 \tag{5}$$

where

$$H_q = \frac{1}{1-q}\left(\int_0^{+\infty} f^q(x)\,\mathrm{d}x - 1\right) \tag{6}$$

is the Tsallis entropy. We note that the same solution is reached if the Tsallis entropy is replaced by using Rényi entropy

$$S_q = \frac{1}{1-q}\log\int_0^{+\infty} f^q(x)\,\mathrm{d}x,$$

since it is a simple monotone transform of the latter. In the limit case $q = 1$, the Tsallis entropy coincides with the Shannon entropy and the maximum entropy solution recovers the canonical exponential distribution

$$f_{1,\beta}(x) = \beta\exp(-\beta x), \quad x \geq 0. \tag{7}$$

The shapes of several q-exponential distributions are shown in Fig. 1 (infinite support case: $q > 1$ and $A_q = +\infty$) and in Fig. 2 (finite support case: $q < 1$ and $A_q = \frac{1}{\beta(1-q)}$); in all cases, parameter β was chosen as $\beta = \frac{1}{2-q}$ so that $f_q(0) = 1$. These figures indicate the high versatility of q-exponentials for modeling probability distributions.

An important and alternate characterization of these distributions has been provided by the superstatistics theory as introduced in Ref. [3] (see also Ref. [4] for related results). In this approach, q-exponential distributions are recognized as scale mixtures of exponential distributions: more precisely, the integral representation

$$f_{q,\frac{1}{q-1}}(x) = \int_0^{+\infty} g_n(\beta)\, f_{1,\beta}(x)\,\mathrm{d}\beta$$

where $g_n(\beta) \propto \beta^{\frac{n}{2}-1}\mathrm{e}^{-\beta}$ is a χ^2_n distribution with $n = \frac{2}{q-1} - 2$ degrees of freedom, allows to describe a q-exponential random variable as a $(q = 1)$ exponential variable with random scale parameter $\beta \sim \chi^2_n$. We note that this approach, in

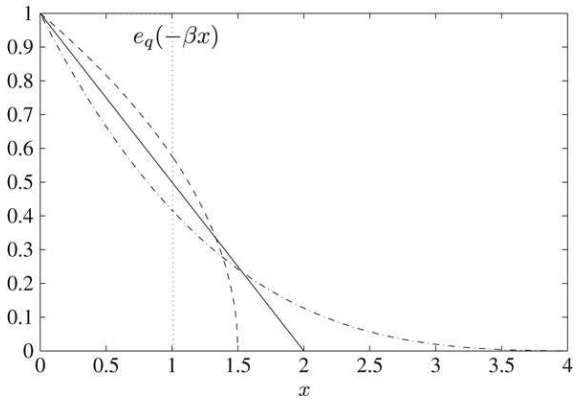


Fig. 2. Finite support q-exponential distributions with $\beta = \frac{1}{2-q}$ and $q = 0$ (solid), $-\infty$ (dot), -1 (dash) and 0.7 (dash-dot).

a physical setting, recovers similar results by Maguire [5] and Harris [6]. This superstatistical approach has found many applications in fluctuation theory: for example, in a recent paper, Briggs and Beck [7] have highlighted the relevance of the superstatistical approach in the modeling of train delays.

Hence, in the one hand, there are two constructions, the maximum q-entropy and the superstatistics one, that exhibit q-exponentials. In the other hand, many observations and experimental results assess the relevance of q-exponentials and their ubiquity. We propose in what follows a third and alternate statistical justification of q-exponential distributions, based on the theory of excess values.

In Section 2, we introduce the "Focus on Excesses" operation, which will reveal as the fundamental operation in our setting. Then, in Section 3 we present the Pickands–Balkema–de Haan result as a possible rationale for the ubiquity of q-exponentials in nature. Furthermore, we underline a stability property of q-exponentials, and show that these results form a Central Limit Theorem for the "Focus on Excesses" operation. In Section 3.2, we draw attention to the Generalized Pareto Distribution (GPD), which is a q-exponential, and underline its applications outside the statistical physics field. Some examples are worked out in Section 3.3. In Section 4, we address the link between the maximum q-entropy and the distributions of excesses. We draw a possible connection and show that the distribution of excesses converges asymptotically to a maximum q-entropy distribution. Finally, we present and discuss in Section 5 the estimation procedures for the parameters of q-exponentials-GPD.

2. The "Focus on Excesses" operation

Let us consider a one-dimensional positive random variable X (ω) that describes the state of a physical system; we denote as $\bar{F}_X$ its survival function

$$\bar{F}_X(x) = 1 - F_X(x) = \Pr\{X > x\}$$

with $F_X(x)$ the distribution function. Let us also consider a positive parameter $u > 0$ and assume that we have access only to the values of X that are larger than u : for example, we may have access to the system through a measurement device that, for a bandwidth or quantification reason, does not "see" any value $X < u$; or we may want to reject all values $X < u$ because we think that they are physically irrelevant. Since parameter u may be large, a wise decision is to shift all these measured values by a factor u : as a consequence, we will in fact record values of the excesses of X over the threshold u, that is values of the *conditional* random variable X_u defined as

$$X_u = X - u | X > u.$$

We denote $X_u = \mathcal{F}_u(X)$ the transform of X into X_u. By Bayes' theorem, the survival function of these excesses writes

$$\bar{F}_{X_u}(x) = \Pr(X > x + u | x > u) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)}, \quad x \geq 0. \tag{8}$$

If X possesses a probability density f_X, then X_u has itself the density

$$f_{X_u}(x) = \frac{f_X(x+u)}{\bar{F}_X(u)}, \quad x \geq 0. \tag{9}$$

The transformation of X into X_u is illustrated in Fig. 3. The inset depicts the excess distribution, i.e. the normalized tail f_{X_u} (9) of the original distribution f_X given in the main part of the figure. For obvious reasons, we call "Focus on Excess" (FoE) operation the transformation of $\bar{F}_X$ into $\bar{F}_{X_u}$, or equivalently of f_X into f_{X_u}. With a slight abuse of notation we also denote $\mathcal{F}_u$ this transformation.

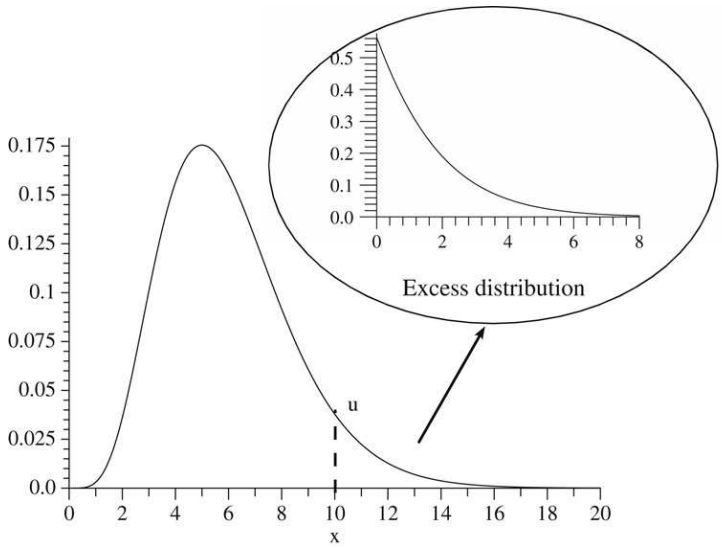


Fig. 3. Illustration of the FoE transformation.

This approach is also known as the "Peaks over Threshold" technique and widely used in fields like climatology and hydrology. A rationale of this method can be found in Ref. [8].

The q-exponential distribution $f_{q,\beta}(x)$ (1) verifies a remarkable *stability* property with respect to this transformation as described in the following theorem (see Ref. [9]).

Theorem 1. *For any $u < A_q$, the FoE of q-exponential remains q-exponential and leaves the tail index q unchanged:*

$$\mathcal{F}_u\left(f_{q,\beta}(x)\right) = f_{q,\beta'}(x)$$

with $\beta' = \frac{\beta}{1-\beta(1-q)u}$.

Proof. By Eq. (9), we know that

$$\mathcal{F}_u\left(e_q(-\beta x)\right) \propto (1-\beta(1-q)(x+u))^{\frac{1}{1-q}} \propto \left(1-\beta'(1-q)x\right)^{\frac{1}{1-q}}, \tag{10}$$

$$= e_q(-\beta' x) \quad \text{for } 0 \le x \le A_q - u \tag{11}$$

where the last term, with $\beta' = \beta/(1-\beta(1-q)u)$, is obtained by factoring $(1-\beta(1-q)u)$. ■

This highlights the particular status of the q-exponential in the FoE transformation. We can also remark that $\beta' = \beta$ in the special exponential ($q = 1$) case: the normalized tail of an exponential distribution is exactly invariant by FoE transformation.

3. The Limit theorem for q-exponentials

3.1. Pickands' theorem

An important theorem in probability as well as in statistical physics is the Central Limit theorem; its extension to the nonextensive context has been the subject of several recent papers [10]. According to this theorem, if a probability density f belongs to the domain of attraction of the Gaussian distribution, and if X_i are centered and independently chosen distributed according to f, then the distribution of

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \tag{12}$$

converges with n to the Gaussian distribution. Moreover, the Gaussian distribution appears as a fixed point in this context since if all X_i are independent and Gaussian in (12) then Z_n is *exactly* Gaussian for any value of n.

Let us denote by $\mathcal{T}_n$ the transform

$$\mathcal{T}_n : f \mapsto f_n$$

where f_n is the distribution of Z_n, i.e. the normalized n-fold convolution of f. Consider now a variable submitted to a series of k successive thresholding operations, with thresholds $u_1, \ldots, u_k$: the resulting excess variable, say Z_u, is

$$Z_u = \mathcal{F}_{u_k} \circ \mathcal{F}_{u_{k-1}} \circ \cdots \mathcal{F}_{u_1}(X)$$

where $\circ$ is the composition of functions and $\mathcal{F}_{u_i}$ as defined above. Of course, the excess variable can also be viewed as resulting from a single thresholding operation, with $u = \sum_{i=1}^{k} u_i$:

$$Z_u = \mathcal{F}_u(X).$$

An analogy appears between $\mathcal{T}_n$ and the FoE transform $\mathcal{F}_u$ which associates the survival functions:

$$\mathcal{F}_u : \bar{F} \mapsto \bar{F}_u,$$

where the real threshold parameter u plays the role of the integer parameter n. Thus, an important question arises at this point: does there exist a limit result, analogous to the CLT, in the context of the FoE transform? Let us first introduce the Fréchet (or heavy-tailed) family of distributions, that will play in extreme statistics the role of the Gaussian domain of attraction in the CLT.

Definition 1. A distribution F belongs to the Fréchet domain if it is a heavy tail distribution, that is if its survival distribution $\bar{F}$ writes

$$\bar{F}(x) = x^{-\frac{1}{\gamma}} l(x), \quad \forall x \geq 0$$

where $l(x)$ is a slowly varying function:

$$\lim_{x \to +\infty} \frac{l(xt)}{l(x)} = 1 \quad \forall t > 0$$

and parameter γ is called the tail index of F.

Pickands' theorem, which can be viewed as the analogue of the CLT in excess statistics, is the following.

Theorem 2. *A necessary and sufficient condition for F to belong to the Fréchet domain is*

$$\lim_{u \to A_q} \|\bar{F}_{X_u}(x) - \bar{F}_{q,\beta(u)}(x)\|_\infty = 0$$

for some function $\beta(u)$ and with

$$q = \frac{2\gamma + 1}{\gamma + 1}. \tag{13}$$

This result means, roughly speaking, that if F is in the domain of attraction of the Fréchet distribution, i.e. has heavy tails with tail index γ, then its FoE transform converges (in the infinite norm sense) to a limit distribution which is nothing but a q-exponential with q given by (13). The accumulation of random variates in the classical CLT, with $n \to \infty$, here corresponds to an increase in the threshold u in the FoE. In the classical CLT, we know that the limit distribution (the Gaussian) once reached remains stable by convolution of Gaussian pdf (addition of Gaussian variables). Accordingly, Theorem 1 shows that q-exponentials are stable by the FoE operation. This means that once the FoE operations have converged to the limit distribution, then it remains stable by further applications of FoE. In other words, q-exponentials are fixed points of the FoE operation.

3.2. Generalized Pareto distributions

In the original Pickands' formulation [11] (see also the similar work by Balkema and De Haan [12]), the name coined for the limit distribution of excesses is Generalized Pareto Distribution (GPD). Its survival function is given by

$$\bar{F}_X(x) = P(X > x) = \left(1 + \frac{\gamma}{\sigma} x\right)_+^{-\frac{1}{\gamma}} \quad \text{for } x \geq 0, \tag{14}$$

where $x_+ = \max(x, 0)$. Its pdf is

$$f_X(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma} x\right)_+^{-\frac{1}{\gamma}-1} \quad \text{for } x \geq 0, \tag{15}$$

where σ and γ are respectively the scale and shape parameters. A location parameter μ can also be specified according to

$$f_X(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma}(x - \mu)\right)_+^{-\frac{1}{\gamma}-1} \quad \text{for } x \geq \mu.$$

Obviously, for $\mu = 0$, the GPD coincides with the q-exponential, with

$$\gamma = -\left(\frac{1-q}{2-q}\right) \quad \text{and} \quad \sigma = \frac{1}{\beta(2-q)}.$$

The interesting point is that the GPD is employed outside the statistical physics field for modeling heavy-tailed distributions. Examples of applications are numerous, ranging from reliability theory [13], traffic in networks [14], hydrology [15,16], climatology [17,18], geophysics [19], materials science [20,21], radar imaging [22] to actuarial sciences [23].

These uses are related to the POT (Peaks over Threshold) method [15,8,24], and the underlying rationale is, of course, the fact that the Balkema–de Haan–Pickands theorem asserts that the distribution of excesses over a threshold often follows approximately a GPD for large values of the threshold.

3.3. Some examples

Let us now detail Pickands' theorem for classical distributions that belong to the Fréchet domain, namely: the alpha-stable, Student and Cauchy distributions.

- **The alpha-stable case.** In the case of an alpha-stable distribution with parameter $0 < \alpha < 1$, we know from Ref. [25, Th.1 p. 448] that the tail behavior of the survival distribution is

  $$\bar{F}_Z(z) \sim \frac{1}{\Gamma(1-\alpha)} z^{-\alpha}$$

  so that the tail index is $\gamma = \frac{1}{\alpha}$ and

  $$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-\alpha}.$$

  We note that the Cauchy distribution ($\alpha = 1$) cannot be deduced from this result.
- **The Student-t case.** A Student-t distribution with $m = 2\nu - 1$ degrees of freedom writes

  $$f_X(x) = A_\nu \left(1 + x^2\right)^{-\nu}, \quad x \in \mathbb{R}$$

  with $A_\nu = \frac{\Gamma(\nu)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\nu - \frac{1}{2}\right)}$. Thus an equivalent of the Student survival distribution is

  $$\bar{F}_Z(z) = \frac{A_\nu}{2\nu - 1} z^{-2\nu+1},$$

  the tail index is $\gamma = (2\nu - 1)^{-1}$ and the excess distribution behaves as

  $$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-2\nu+1}.$$
- **The Cauchy case.** As a special case, the Cauchy distribution corresponds to $m = 1$ degrees of freedom and

  $$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-1}. \tag{16}$$

In nonextensive statistics, the Student-t distributions described above are well known under the name of q-Gaussian distributions ($q > 1$). What is shown here is that the excess distribution of a q-Gaussian distribution with nonextensivity parameter q converges to a q-exponential distribution with nonextensivity parameter

$$q' = \frac{1+q}{2}.$$

We remark that in the limit case where $q \to 1$, the q-Gaussian distribution converges to the classical Gaussian distribution. In this case we also obtain that $q' \to 1$, which means that the excess distribution of a Gaussian converges to the exponential distribution. Indeed, Pickands' theorem extends to the Gumbel family of distributions characterized by an exponential tail: the associated excess pdf is simply the distribution given by (7), which is the limit case of the q-exponential (1) for $q \to 1$.

3.4. Numerical illustration

As a numerical illustration, we present the Cauchy case

$$f_X(x) = \frac{2}{\pi\left(1 + x^2\right)}, \quad x \geq 0.$$

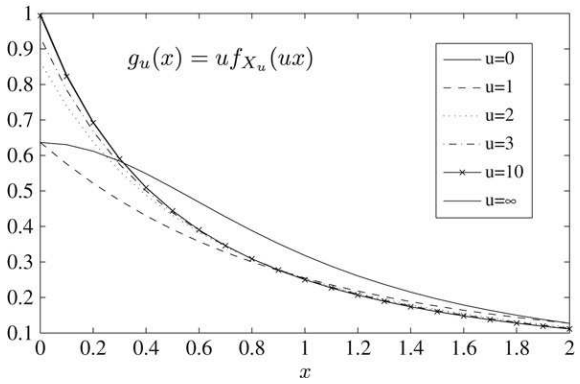


Fig. 4. Illustration of Pickands' theorem in the Cauchy case. The figure presents the normalized excess pdfs $g_u(x)$, for several values of the threshold u. It shows that the normalized pdf quickly converges, as u increases, to the limit distribution $1/(1+x)^2$.

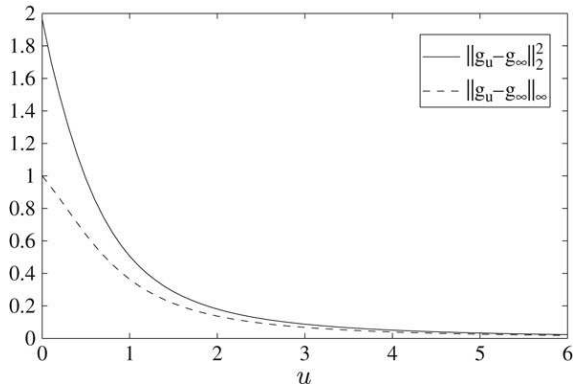


Fig. 5. Rate of convergence to the limit distribution. The rate of convergence is illustrated by the quadratic and infinite norms of the approximation error $g_u - g_\infty$, as a function of the threshold u.

The survival function of the excess variable X_u is

$$\bar{F}_{X_u}(x) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)} = \frac{1 - \frac{2}{\pi}\arctan(x+u)}{1 - \frac{2}{\pi}\arctan(u)}.$$

Using the fact that $\arctan(t) \approx \pi/2 - 1/t$, for $t \gg 1$, we have

$$\bar{F}_{X_u}(x) \approx \frac{1}{u+x}\Big/\frac{1}{u} = \left(1 + \frac{x}{u}\right)^{-1}$$

which is (16) again. It is the survival function of a q-exponential with entropic index $q = 3/2$.

The excess density can be explicited as

$$f_{X_u}(x)\,(x) = \frac{2}{\pi}\frac{1}{\left(1 - \frac{2}{\pi}\arctan(u)\right)\left(1 + (x+u)^2\right)}.$$

The convergence of the survival distribution to a Generalized Pareto Distribution with $\gamma = 1$ as in (16) is illustrated in Fig. 4. In this figure, we have plotted the normalized pdfs $x \mapsto uf_{X_u}(ux) = g_u(x)$, corresponding to the normalized random variable $Y_u = X_u/u$, for several values of the threshold u. The figure shows the convergence to the GPD limit $g_\infty(x) = \frac{1}{(1+x)^2}$. We note that Pickands' theorem only ensures convergence of the distribution functions, but the Cauchy case is smooth enough to observe a convergence of the normalized pdf as well. We remark on Fig. 4 that small values of parameter u (typically $u = 3$) ensure an already accurate approximation of f_{X_∞}. Rate of convergence to the limit distribution are examined in Fig. 5, where we report the quadratic and infinite norms of the approximation error, as a function of u:

$$\|g_u - g_\infty\|_\infty \quad \text{and} \quad \|g_u - g_\infty\|_2^2.$$

Note that the rate of convergence of the GPD approximation in Pickands' theorem is rigorously quantified in Ref. [26].

4. A connection with maximum q-entropy

Since there exists an entropic proof of the classical central limit theorem, it is appealing to look for relationships between Pickands' result and the maximum entropy principle. We show here that a distribution of the excesses over a threshold converges, as this threshold goes to infinity, to a maximum q-entropy solution. Details of similar derivations are given in Ref. [27].

Proposition 3 gives a simplified form of a q-exponential distribution function as the solution of a maximum q-norm problem. Then, we state in Theorem 4 that the 1-norm, q-norm, and q-expectation of a suitably normalized version of the excess variable (in the Fréchet domain) converges asymptotically to constant values. Finally, with an appropriate choice of q, this shows that the distribution of excesses is necessarily, asymptotically, the q-exponential solution of the maximum q-norm problem (or equivalently of Tsallis' q-entropy maximization).

Proposition 3. *With $\sqrt{2} > q > 1$, consider the class of functions*

$$\mathcal{F} = \left\{ G : \mathbb{R}^+ \to \mathbb{R}; \int_0^{+\infty} zG(z)\,\mathrm{d}z < \infty \right\}.$$

The function

$$G_*(z) = (1+z)^{\frac{1}{1-q}+1} \tag{17}$$

is the unique solution of the following maximum norm problem:

$$\max_{G\in\mathcal{F}} \|G\|_q^q = \max_{G\in\mathcal{F}} \int_0^{+\infty} G(z)^q\,\mathrm{d}z$$

such that

$$\int_0^{+\infty} zG(z)^q\,\mathrm{d}z = \frac{(q-1)^2}{(q^2-q-1)(q^2-2)} \quad \text{and} \quad \int_0^{+\infty} G(z)\,\mathrm{d}z = \frac{1-q}{2q-3}. \tag{18}$$

Moreover, the corresponding maximum norm is

$$\|G_*\|_q^q = \int_0^{+\infty} G_*(z)^q\,\mathrm{d}z = \frac{1-q}{q^2-q-1}.$$

Let us now consider a random variable X that belongs to the Fréchet domain with tail index $\gamma < \frac{1}{2}$, and a normalized version $Y_u = X_u/g(u)$ of its excesses, where function $g(u) \sim u$, as $u \to +\infty$. Then, we have the following results.

Theorem 4. *With $\gamma < q < 1$, as $u \to +\infty$, the following asymptotics hold*

$$\|\bar{F}_{Y_u}\|_q^q \sim \frac{1}{\frac{q}{\gamma}-1}$$

with asymptotical q-expectation

$$\int_0^{+\infty} z\bar{F}_{Y_u}(z)^q\,\mathrm{d}z \sim \frac{1}{\left(1-\frac{q}{\gamma}\right)\left(2-\frac{q}{\gamma}\right)}$$

and asymptotical 1-norm

$$\int_0^{+\infty} \bar{F}_{Y_u}(z)\,\mathrm{d}z = \frac{1}{\frac{1}{\gamma}-1}.$$

As a consequence, with the choice of

$$q = \frac{1+2\gamma}{1+\gamma} \quad \text{or } \gamma = \frac{1-q}{q-2}$$

we obtain that the excess distribution of any suitably normalized random variable $Y = X/g(u)$, where X belongs to the Fréchet domain and $g(u) \sim u$, has asymptotically same 1-norm, q-norm and q-expectation as the q-exponential distribution (17). Since the maximum norm problem (18) has a unique solution, we deduce that the maximum q-norm distribution G_* and the distribution of excesses F coincide asymptotically as $u \to +\infty$. The physical significance behind this result is that it enables to connect the distribution of the excesses over a threshold and a maximum q-entropy construction, and thus add a

new motivation for the use of Tsallis entropy: q-exponentials, as the Tsallis entropy maximizers, can arise in many contexts in which the system at hand or the measurement device introduces some threshold.

Moreover, function G_* (17) can now be identified as the survival function associated to the q-exponential density

$$g_*(x) \propto (1+x)^{\frac{1}{1-q}},$$

which is itself the solution of a maximum q-norm problem (5) similar to (18). Hence, the asymptotic density of excesses (if it exists) is a q-exponential density. The addition of a scale parameter $\beta(q-1)$ simply enables to adjust the mean or variance of the distribution.

5. Parameter estimation

The statistical analysis of heavy-tailed data requires some caution [28]. A recent reference for this topic is Ref. [29].

An important issue for applications of q-exponential-GPD and the assessment of their role in real situations is the estimation of their parameters. As far as the tail exponent $\alpha = 1/\gamma$ is concerned, several methods exist. The corresponding, and relevant problem, in the nonextensive framework is the estimation of parameter q.

The Hill estimator. The Hill estimator [30] is given by

$$\hat{\gamma}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\log X_{(n-i+1)} - \log X_{(n-k)}\right)$$

where $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ are the order statistics of the sample of size n, and k is a smoothing parameter. This estimator, which can be derived by several rationale (including maximum likelihood) is valid for $\gamma > 0$, i.e. $q > 1$.

Pickands' estimator. Another, more general estimator, is the Pickands' estimator, which is valid for all γ and given by

$$\hat{\gamma}(k) = \frac{1}{\log 2}\log\frac{X_{(n-k+1)} - X_{(n-2k+1)}}{X_{(n-2k+1)} - X_{(n-4k+1)}}.$$

Both Hill and Pickands' estimators are consistent and asymptotically normal [31,29].

Maximum Likelihood. The log-likelihood for a set of n independent and identically distributed observations X_i, distributed according to a GPD (γ, σ) given by (15), is:

$$\log p_n(X_1, X_2, \ldots, X_n) = -n\log\sigma - \left(\frac{1}{\gamma}+1\right)\sum_{i=1}^{n}\log\left(1+\frac{\gamma}{\sigma}X_i\right).$$

Then, the Maximum Likelihood procedure gives two equations for the unknown parameters. These relations are implicit but can be solved numerically. We obtain

$$\sigma = \frac{1+\gamma}{n}\sum_{i=1}^{n}\frac{X_i}{1+\frac{\gamma}{\sigma}X_i}$$

and

$$\gamma = \frac{1}{n}\sum_{i=1}^{n}\log\left(1+\frac{\gamma}{\sigma}X_i\right).$$

More details on Maximum Likelihood estimation of GPD parameters can be found in Refs. [32,33]. The Maximum Likelihood for q-exponentials has also been discussed recently in Ref. [34]. It shall be mentioned that the ML does not exist for $\gamma < -1$, which corresponds to the case $q > 2$, where $e_q(x)$ is not normalizable.

Method of moments. When both the mean m_X and variance s_X^2 exist, that is for $\gamma < 0.5$ $(q > 0.5)$, one can apply a *method of moments* (MOM) since there is a simple relationship between these moments and the parameters γ and σ, namely

$$\sigma = 0.5 m_X\left(1 + (m_X/s_X)^2\right) \tag{19}$$

and

$$\gamma = 0.5\left(1 - (m_X/s_X)^2\right) \tag{20}$$

then it suffices to use some estimates of the mean and standard deviation in (19) and (20) to obtain a (rough) estimate of the GPD parameters.

Conditional mean exceedance. Variations on the method of moments include the so-called "conditional mean exceedance method" which relies on the analysis of the plot of the mean of excesses over u, which is linear according to

$$E[X-u|X>u] = \frac{\sigma+\gamma u}{1-\gamma},$$

with slope $\gamma/(1-\gamma)$ and intercept $\sigma/(1-\gamma)$. For several values of the threshold u, and therefore estimates of the mean of excesses, a least-squares procedure can be used in order to identify the parameters of the linear model.

Probability Weighted Moments. An alternative method to the MOM, known as "Probability Weighted Moments" (PWM) was introduced by Refs. [9,35]. The paper [9] also provides an interesting comparison of ML, MOM and PWM estimates. This procedure relies on the definition of a "weighted moment":

$$m_p = E\left[X\bar{F}_X(X)^p\right]$$

where $\bar{F}_X$ is the survival function of the GPD given in (14). For $\gamma < 1$ ($q > 0$), we readily obtain that

$$m_p = \frac{\sigma}{(p+1)(p+1-\gamma)}.$$

Then, two weighted moments are sufficient to exhibit the values of the parameters γ and σ; e.g. with m_1 and m_0, one has

$$\gamma = \frac{4m_1 - m_0}{2m_1 - m_0} \quad \text{and} \quad \beta = \frac{2m_1 m_0}{m_0 - 2m_1}.$$

With more moments, estimated classically, one can look for a least-squares solution. PWM is reported to compete with the Maximum Likelihood.

Synthesis and software. Another method is the Elemental Percentile Method, discussed in Ref. [36]. A synthesis of estimation methods of GPD parameters can be found in Ref. [33, chapter 20, pages 614-620] and in Ref. [37].

It shall be mentioned that statistical softwares include estimation methods for GPD-q-exponential parameters. An example is the "Vector Generalized Additive Models" package [38,39]. Another example is the statistical toolbox of the highly employed Matlab™ software, which includes fitting tools for the GPD since its 5.1 version (2006).

6. Conclusion

In this paper, we have proposed a possible rationale for understanding the ubiquity of q-exponential distributions in nature. The point is that, given a heavy-tailed distributed system, q-exponentials can occur as soon as the measurement device, or the system at hand, involves some threshold. This is the essence of Pickands' result.

Furthermore, the q-exponentials are stable by thresholding. This shows that q-exponentials are the limit distributions in a Central Limit like theorem, where the underlying operation is the "Focus on Excesses" operation we have introduced.

We have also underlined that q-exponentials can also be recognized as Generalized Pareto Distributions which are of importance in many applications outside the statistical physics field.

We have drawn a connection between excess distributions and the maximum q-entropy, by showing that a distribution of excesses converges to a maximum q-entropy solution.

Finally, we have presented some procedures for the crucial problem of estimation of q-exponential parameters. We have also provided a large number of bibliographic entries on these different topics.

In our view, the rationale proposed here is complementary to the classical maximum q-entropy approach, and to the Superstatistics approach. We hope that it will prompt some new viewpoints in physical applications.

References

[1] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, Journal of Statistical Physics 52 (1988) 479–487.
[2] H. Hilhorst, G. Schehr, A note on q-Gaussians and non-Gaussians in statistical mechanics, Journal of Statistical Mechanics: Theory and Experiment (P06003).
[3] C. Beck, E.G.D. Cohen, Superstatistics, Physica A 322 (2003) 267.
[4] G. Wilk, Z. Wlodarczyk, Interpretation of the nonextensivity parameter q in some applications of Tsallis statistics and Lévy distributions, Physical Reviews Letters 84 (2000) 2770.
[5] B.A. Maguire, E.S. Pearson, H.A. Wynn, The time intervals between industrial accidents, Biometrika 39 (1952) 168–180.
[6] C.M. Harris, The Pareto distribution as a queue service discipline, Operations Research 16 (1968) 307–313.
[7] K. Briggs, C. Beck, Modelling train delays with q-exponential functions, Physica A 378 (2007) 498.
[8] R. Leadbetter, On a basis for 'Peaks over Threshold' modeling, Statistics and Probability Letters 12 (1991) 357–362.
[9] J.R.M. Hosking, J.R. Wallis, Parameter and quantile estimation for the Generalized Pareto Distribution, Technometrics 29 (1987) 339–349.
[10] S. Umarov, C. Tsallis, On multivariate generalizations of the q-central limit theorem consistent with nonextensive statistical mechanics, in: S. Abe, H. Herrmann, P. Quarati, A. Rapisarda, C. Tsallis (Eds.), Complexity, Metastability and Nonextensivity, Vol. 965, American Institute of Physics Conference Proceedings, New York, 2007, p. 34.
[11] J. Pickands, Statistical inference using extreme order statistics, The Annals of Statistics 3 (1975) 119–131.
[12] A.A. Balkema, L. de Haan, Residual life time at great age, The Annals of Probability 2 (1974) 792–804.
[13] M. Asadi, I. Bayramoglu, The mean residual life function of a k-out-of-n structure at the system level, IEEE Transactions on Reliability 55 (2006) 314–318.
[14] P.-H. Hsieh, An exploratory first step in teletraffic data modeling: Evaluation of long-run performance of parameter estimators, Computational Statistics & Data Analysis 40 (2002) 263–283.
[15] M.A.J. Van Montfort, J.V. Witter, Generalized Pareto Distribution applied to rainfall depths, Hydrological Sciences Journal 31 (1986) 151–162.
[16] S. Coles, L.R. Pericchi, S. Sisson, A fully probabilistic approach to extreme rainfall modeling, Journal of Hydrology 273 (2003) 35–50.
[17] S. Coles, Casson, Extreme value modelling of hurricane wind speeds, Structural Safety 20 (1998) 283–296.
[18] I. Harris, Generalised Pareto methods for wind extremes, Useful tool or mathematical mirage? Journal of Wind Engineering and Industrial Aerodynamics 93 (2005) 341–360.

[19] V.F. Pisarenko, D. Sornette, Characterization of the frequency of extreme events by the Generalized Pareto Distribution, Pure and Applied Geophysics 160 (2003) 2343.
[20] G. Shi, H.V. Atkinson, C.M. Sellars, C.W. Anderson, Application of the Generalized Pareto Distribution to the estimation of the size of the maximum inclusion in clean steels, Acta Materialia 47 (1999) 1455–1468.
[21] D. Fullwood, B. Adams, S. Kalidindi, Generalized Pareto front methods applied to second-order material property closures, Computational Materials Science 38 (2007) 788–799.
[22] B. LaCour, Statistical characterization of active sonar reverberation using extreme value theory, IEEE Journal of Oceanic Engineering 29 (2004) 310–316.
[23] A.C. Cebrian, M. Denuit, P. Lambert, Generalized Pareto fit to the society of actuaries' large claims database, North American Actuarial Journal 7 (2003) 18–36.
[24] J.A. Lechner, E. Simiu, N.A. Heckert, Assessment of 'Peaks over Threshold' methods for estimating extreme value distribution tails, Structural Safety 12 (1993) 305–314.
[25] W. Feller, An introduction to Probability theory and its applications, vol. 2, John Wiley and Sons, 1971.
[26] R. Worms, Vitesse de convergence de l'approximation de Pareto généralisée de la loi des excès – Rate of convergence for the generalized Pareto approximation of the excesses, Comptes Rendus de l'Academie des Sciences - Series I - Mathematics 333 (2001) 65–70.
[27] J.F. Bercher, C. Vignat, An entropic view of Pickands' theorem, in: International Symposium on Information Theory ISIT2008, Toronto, 2008. URL http://arxiv.org/abs/0802.3110.
[28] A. Clauset, C.R. Shalizi, M.E.J. Newman, Power-law distributions in empirical data, 0706.1062.
[29] N. Markovich, Nonparametric Analysis of Univariate Heavy-Tailed Data, in: Series in Probability and Statistics, Wiley, 2007.
[30] B. Hill, A simple general approach to inference about the tail of a distribution, Annals of Statistics 3 (1975) 1163–1174.
[31] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, 1st Edition, Springer, 2008.
[32] S.D. Grimshaw, Computing maximum likelihood estimates for the Generalized Pareto distribution, Technometrics 35 (1993) 185–191.
[33] N.L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, 2nd Edition, Wiley-Interscience, 1994.
[34] C.R. Shalizi, Maximum likelihood estimation for *q*-exponential (Tsallis) distributions, math/0701854.
[35] J. Hosking, Some theory and practical uses of trimmed l-moments, Journal of Statistical Planning and Inference 137 (2007) 3024–3039.
[36] E. Castillo, A.S. Hadi, Fitting the generalized pareto distribution to data, Journal of the. American Statistical Association 92 (1997) 1609–1620.
[37] J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, Statistics of Extremes: Theory and Applications, 1st Edition, Wiley, 2004.
[38] T.W. Yee, C.J. Wild, Vector Generalized Additive Models, Journal of the Royal Statistical Society, Series B, Methodological 58 (1996) 481–493.
[39] T.W. Yee, The VGAM package (2007). URL http://www.stat.auckland.ac.nz/~yee/VGAM/.

8.5 J.-F. Bercher, "On some entropy functionals derived from Rényi information divergence," *Information Sciences*, vol. 178, pp. 2489–2506, Jun. 2008.

Available online at www.sciencedirect.com
ScienceDirect
Information Sciences 178 (2008) 2489–2506
INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL
www.elsevier.com/locate/ins

On some entropy functionals derived from Rényi information divergence

J.-F. Bercher [1]

Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette cedex, France

Received 9 November 2007; accepted 11 February 2008

Abstract

We consider the maximum entropy problems associated with Rényi Q-entropy, subject to two kinds of constraints on expected values. The constraints considered are a constraint on the standard expectation, and a constraint on the generalized expectation as encountered in nonextensive statistics. The optimum maximum entropy probability distributions, which can exhibit a power-law behaviour, are derived and characterized.

The Rényi entropy of the optimum distributions can be viewed as a function of the constraint. This defines two families of entropy functionals in the space of possible expected values. General properties of these functionals, including nonnegativity, minimum, convexity, are documented. Their relationships as well as numerical aspects are also discussed. Finally, we work out some specific cases for the reference measure $Q(x)$ and recover in a limit case some well-known entropies.

Keywords: Rényi entropy; Rényi divergences; Maximum entropy principle; Nonextensivity; Tsallis distributions

1. Introduction

Consider two univariate continuous probability distributions with densities P and Q with respect to the Lebesgue measure. The Rényi information divergence introduced in [32] has the form

$$D_\alpha(P||Q) = -H_Q^{(\alpha)}(P) = \frac{1}{\alpha - 1} \log \int_{\mathscr{D}} P(x)^\alpha Q(x)^{1-\alpha} \, \mathrm{d}x, \tag{1}$$

where α is a positive real and $\mathscr{D}$ the domain of definition of the integral. In the discrete case, the continuous sum is replaced by a discrete one which extends on a subset $\mathscr{D}$ of integers. The opposite $H_Q^{(\alpha)}(P)$ of the Rényi information divergence can be viewed as a Rényi entropy relative to the reference measure Q, and can be called Q-entropy. By L'Hospital's rule, Kullback divergence is recovered in the limit $\alpha \to 1$.

Applications and areas of interest in Rényi entropy are plentiful: communication and coding theory [10], data mining, detection, segmentation, classification [29,5], hypothesis testing [23], characterization of signals

E-mail address: jf.bercher@esiee.fr
[1] On sabbatical leave from ESIEE-Paris, France.

0020-0255/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2008.02.003

and sequences [38,19], signal processing [5,3], image matching and registration [29,15]. Connection with the log-likelihood has been outlined in [33], where is also defined a measure of the intrinsic shape of a distribution which can serve as a measure of tail heaviness [27]. Rényi entropies for large families of univariate and bivariate distributions are given in [25,26]. Divergence measures based on entropy functions can be used in the process of inference [12], in clustering or partitioning problems [22,2,7].

Rényi entropy also plays a central role in the theory of multifractals, see for instance [18,4]. In statistical physics, following Tsallis proposal [34,35] of another entropy (which is simply related to Rényi entropy), there has been a high interest on these alternative entropies and the development of a community in 'nonextensive thermostatistics'. Indeed, the associated maximum entropy distributions exhibit a power-law behaviour, with a remarkable agreement with experimental data, see for instance [6,35] and references therein. These optimum distributions, called Tsallis distributions, are similar to Generalized Pareto Distributions, which also have an high interest in other fields, namely reliability theory [1], climatology [24], radar imaging [21] or actuarial sciences [8].

Jaynes' maximum entropy principle [16,17] suggests that the least biased probability distribution that describes a partially-known system is the probability distribution with maximum entropy compatible with all the available prior information. When prior information is available in the form of constraints on expected values, the maximum entropy method amounts to minimize Kullback information divergence $D(P\|Q)$ (or equivalently maximizing Shannon Q-entropy) subject to normalization and an observation constraints. In the case of a single constraint on the mean of the distribution, say $E_P[X] = m$, the minimum of Kullback information in the set of all probability distributions with expectation m is of course a function of m, denoted $\mathscr{F}(m)$ as follows:

$$\mathscr{F}(m) = \begin{cases} \min\limits_{P} \; D(P||Q) \\ \text{s.t. } m = E_P[X] \\ \text{and } \int_{\mathscr{D}} P(x)\,\mathrm{d}x = 1. \end{cases} \tag{2}$$

It is a 'contracted' version of Shannon Q-entropy and is called a level-1 entropy functional, or rate function, in the theory of large deviations, e.g. [11]. The maximum entropy method is a widely and successful method extensively used in a large variety of problems and contexts.

We focus here on solutions and properties of maximum entropy problems analog to (2) for the Rényi information divergence (1), and on the associated entropy functionals. The maximum Rényi–Tsallis entropy distribution, with its power-law behaviour, is at the heart of nonextensive statistics, but have also be considered in [13,14]. In nonextensive statistics, one still consider the usual classical mean constraint, but also a 'generalized' α-expectation constraint. This 'generalized' α-expectation is in fact the expectation with respect to the distribution

$$P^*(x) = \frac{P(x)^{\alpha} Q(x)^{1-\alpha}}{\int_{\mathscr{D}} P(x)^{\alpha} Q(x)^{1-\alpha}\,\mathrm{d}x}, \tag{3}$$

that is a weighted geometric mean of P and Q. It is nothing else but the 'escort' or zooming distribution of nonextensive statistics [35] and multifractals. Of course, with $\alpha = 1$, the escort distribution P^* reduces to P and the generalized mean $E_{P^*}[X]$ reduces to the standard one.

Therefore, the maximum entropy problems associated to Rényi information divergence (1), subject to normalization and to a classical (C) or generalized (G) mean constraint states as

$$\mathscr{F}_{\alpha}^{(C \text{ resp. } G)}(m) = \begin{cases} \min\limits_{P} \; D_{\alpha}(P||Q) \\ \text{s.t. } (C)m = E_P[X] \\ \text{or } (G)m = E_{P^*}[X] \\ \text{and } \int_{\mathscr{D}} P(x)\,\mathrm{d}x = 1, \end{cases} \tag{4}$$

where $\mathscr{F}_{\alpha}^{(C)}(m)$ and $\mathscr{F}_{\alpha}^{(G)}(m)$ are the level-one entropy functionals associated to Rényi Q-entropy for the classical an generalized constraints respectively. Since Rényi entropy reduces to Shannon's for $\alpha = 1$, functionals $\mathscr{F}_{\alpha}^{(\cdot)}(m)$ will reduce to $\mathscr{F}(m)$ when $\alpha \to 1$.

Hence, in this paper, we consider the forms and properties of maximum entropy solutions associated to Rényi Q-entropy, subject to two kind of constraints, as explained above. The value of the maximum entropy problems at the optimum define entropy functionals $\mathscr{F}_\alpha^{(\cdot)}(m)$ associated to each choice of reference Q, and indexed by the parameter α. The introduction of the reference measure Q, and therefore the definition of functionals $\mathscr{F}_\alpha^{(\cdot)}(m)$ is, to the best of our knowledge, new in this setting. In Section 2, the exact form of the probability distributions P that realize the minimum of the Rényi information divergence in the right side of (4) are first derived. Then we give some properties of these distributions and of their partition functions. We show that the entropy functionals $\mathscr{F}_\alpha^{(\cdot)}(m)$ are simply linked to these partition functions. General properties of the entropy functionals, including nonnegativity, convexity, are established. We also indicate how the problems (4) can be tackled numerically, for specific values of the constraints, even thouh the maximum entropy distributions exhibit implicit relationships. A divergence in the object space, that reduces to a Bregman divergence for $\alpha \to 1$ is defined. These results are illustrated in Section 3 where we study four special cases of reference Q, and characterize the associated entropy functionals. It is then shown that some well-known entropies are recovered.

2. The minimum of Rényi divergence

Let us define by

$$P_v(x) = \frac{[1+\gamma(x-\bar{x})]^v}{Z_v(\gamma,\bar{x})} Q(x), \tag{5}$$

a probability density function on a subset $\mathscr{D}$ of $\mathbb{R}$, where $\mathscr{D}$ ensure that the numerator of (5) is always nonnegative and its integral finite. The normalization $Z_v(\gamma,\bar{x})$ is the partition function defined by

$$Z_v(\gamma,\bar{x}) = \int_{\mathscr{D}} [1+\gamma(x-\bar{x})]^v Q(x)\,\mathrm{d}x. \tag{6}$$

The density P_v depends of three parameters: the exponent v which can be considered as a shape parameter, a scale parameter γ and a location parameter $\bar{x}$. But these parameters can be also be linked. For instance, $\bar{x}$ might be a function of v and γ. When non ambiguous, we may also denote by $E_v[X]$ the statistical mean with respect to $P_v(x)$.

With these notations, we have the following result.

Theorem 1. *(C) The distribution $P_C(x)$ in the family (5) with $v = \xi = \frac{1}{\alpha-1}$ and $\bar{x} = E_P[X] = E_\xi[X]$, has the minimum Rényi divergence to Q*

$$D_\alpha(P||Q) \geqslant D_\alpha(P_C||Q) \tag{7}$$

for all probability distributions $P(x)$ absolutely continuous with respect to $P_C(x)$ with a given (classical) expectation $\bar{x}$.

(G) The distribution $P_G(x)$ in the family (5) with $v = -\xi = \frac{1}{1-\alpha}$ and $\bar{x} = E_{P_G^*}[X] = E_{-(\xi+1)}[X]$, has the minimum Rényi divergence to Q

$$D_\alpha(P||Q) \geqslant D_\alpha(P_G||Q) \tag{8}$$

for all probability distributions $P(x)$ absolutely continuous with respect to $P_G(x)$ with a given generalized expectation $\bar{x}$.

Corollary 2. *The solution to the minimization of Rényi divergence in (4) is as given in Theorem 1 for the particular values γ^* of γ such that $\bar{x} = m$.*

It is important to emphasize that $\bar{x}$ is here a statistical mean, and not the constraint m, and as such a function of γ.

Proof. See Appendix A. □

Remark 3. When α tends to 1, $|v|$ tends to $+\infty$. Let us introduce $\tilde{\gamma}$ such that $\gamma = \tilde{\gamma}/v$. Then

$$P_v(x) = \mathrm{e}^{v\log\left[1+\frac{\tilde{\gamma}}{v}(x-\bar{x})\right]-\log Z_v(\tilde{\gamma},\bar{x})}Q(x), \tag{9}$$

and

$$\lim_{|v|\to+\infty} P_v(x) = \mathrm{e}^{\tilde{\gamma}(x-\bar{x})-\log Z_v(\tilde{\gamma},\bar{x})}Q(x), \tag{10}$$

that is the standard exponential, which is the well-known solution of the minimisation of Kullback–Leibler divergence subject to a constraint on an expected value [20, Theorem 2.1, p. 38]. In this case, the log-partition function becomes

$$\lim_{|v|\to+\infty} \log Z_v(\tilde{\gamma},\bar{x}) = \tilde{\gamma}\bar{x} - \log\int_{\mathscr{D}} \mathrm{e}^{\tilde{\gamma}x}Q(x)\,\mathrm{d}x. \tag{11}$$

Properties of entropy functionals $\mathscr{F}_\alpha^{(C)}(m)$ and $\mathscr{F}_\alpha^{(G)}(m)$ are of course linked to the properties of the optimum distribution (5) and its partition function (6). In Property 4, we characterize partition functions of successive exponents, which enables to derive the expression of the Rényi entropy associated to the optimum distribution. In Proposition 6, we give the expression of the derivative of the partition function with respect to γ. Since the optimum distribution (5) is 'self-referential' (because it depends of its mean, which gives an implicit relation), direct determination of its parameters is difficult. It could rely on tabulation or on iterative techniques [36], that still suppose that the solution is an attractive fixed point. We define in Proposition 9 two functionals whose maximization provide the γ parameter of the optimum distributions associated to the classical and generalized mean constraint. Then general properties of nonnegativity, minimum, convexity are then given in Proposition 11. We also show that the two entropy optimization problems are related and that functionals $\mathscr{F}_\alpha^{(\cdot)}(m)$ obey a special symmetry. Finally, we define a divergence in the space of possible means.

Property 4. *Partition functions of successive exponents are linked by*

$$Z_{v+1}(\gamma,\bar{x}) = E_{v+1-k}[(\gamma(x-\bar{x})+1)^k]Z_{v+1-k}(\gamma,\bar{x}). \tag{12}$$

An interesting particular case is for $k = 1$:

$$Z_{v+1}(\gamma,\bar{x}) = E_v[\gamma(x-\bar{x})+1]Z_v(\gamma,\bar{x}). \tag{13}$$

This is easily checked by direct calculation. As a direct consequence, we may also observe that $Z_{v+1}(\gamma,\bar{x}) = Z_v(\gamma,\bar{x})$ if and only if $\bar{x} = E_v[x]$. When $\bar{x}$ is a fixed parameter m, this will be only true for a special value γ^* such that $E_v[x] = m$.

Now, using (13) in Property 4, it is possible to give the expression of the Rényi divergence associated to the distribution (5) and in particular to the solutions P_C and P_G of problems (4):

Property 5. *The Rényi information divergence associated to the optimum distributions (5) in Theorem 1 is* (*C*) $D_\alpha(P||Q) = -\log Z_\xi(\gamma,\bar{x}) = -\log Z_{\xi+1}(\gamma,\bar{x})$, *and* (*G*) $D_\alpha(P||Q) = -\log Z_{-\xi}(\gamma,\bar{x}) = -\log Z_{-(\xi+1)}(\gamma,\bar{x})$.

Proof. The Rényi entropy associated to (5) writes

$$D_\alpha(P||Q) = \frac{1}{\alpha-1}\log\int P(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x = \frac{1}{\alpha-1}\log\int(1+\gamma(x-\bar{x}))^{\alpha v}Q(x)\,\mathrm{d}x - \frac{\alpha}{\alpha-1}\log Z_v(\gamma,\bar{x}),$$

that simply reduces to

$$D_\alpha(P||Q) = \frac{1}{\alpha-1}\log Z_{\alpha v}(\gamma,\bar{x}) - \frac{\alpha}{\alpha-1}\log Z_v(\gamma,\bar{x}).$$

(*C*) In one hand, if $v = \xi = \frac{1}{\alpha-1}$, then $\alpha v = \frac{\alpha}{\alpha-1} = \xi+1$, and $D_\alpha(P||Q) = \frac{1}{\alpha-1}\log Z_{\xi+1}(\gamma,\bar{x}) - \frac{\alpha}{\alpha-1}\log Z_\xi(\gamma,\bar{x})$ Therefore, when $\bar{x} = E_\xi[x]$, then (13) gives $Z_{\xi+1}(\gamma,\bar{x}) = Z_\xi(\gamma,\bar{x})$, and it simply remains

$$D_\alpha(P||Q) = -\log Z_\xi(\gamma,\bar{x}) = -\log Z_{\xi+1}(\gamma,\bar{x}).$$

(G) In the other hand, if $v = -\xi = \frac{1}{1-\alpha}$, then $\alpha v = \frac{\alpha}{1-\alpha} = -\xi - 1$, and $D_\alpha(P||Q) = \frac{1}{\alpha-1}\log Z_{-(\xi+1)}(\gamma,\bar{x}) - \frac{\alpha}{\alpha-1}\log Z_{-\xi}(\gamma,\bar{x})$. When $\bar{x} = E_{-(\xi+1)}[x]$, we have $Z_{-\xi}(\gamma,\bar{x}) = Z_{-(\xi+1)}(\gamma,\bar{x})$ according to (13) and it remains

$$D_\alpha(P||Q) = -\log Z_{-\xi}(\gamma,\bar{x}) = -\log Z_{-(\xi+1)}(\gamma,\bar{x}). \quad \square$$

Since the Rényi information divergence of distributions (5) is simply the log-partition function, it will be useful to examine the behaviour of the partition function with respect to the parameter γ. Hence, the following proposition gives the expression of the derivative of the partition function.

Proposition 6. *For the partition function (6) with domain of definition $\mathscr{D}$, the derivative with respect to γ of the partition function with characteristic exponent v is given by*

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} Z_v(\gamma,\bar{x}) = v\left(E_{v-1}[x-\bar{x}] - \gamma\frac{\mathrm{d}\bar{x}}{\mathrm{d}\gamma}\right) Z_{v-1}(\gamma,\bar{x}). \tag{14}$$

if (a) the domain $\mathscr{D}$ does not depend of γ, or (b) on subsets of γ such that the domain increment $\delta\mathscr{D}$ associated to the variation $\delta\gamma$ remains empty, or (c) for $v > 0$ in the continuous case or $v > 1$ in the discrete case.

Proof. See Appendix B. $\square$

Using this proposition on the derivative of the partition function and Property 4 on the link between partitions functions of succesive exponents, we readily have

Property 7. *If $\bar{x} = E_{v-1}[X]$, then, with the same conditions as in Proposition 6:*

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\log Z_v(\gamma,\bar{x}) = -\gamma v\frac{\mathrm{d}\bar{x}}{\mathrm{d}\gamma} \tag{15}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\bar{x}}\log Z_v(\gamma,\bar{x}) = -\gamma v. \tag{16}$$

This is immediately checked using (13) and (14) with $\bar{x} = E_{v-1}[X]$. It is now interesting to consider the special case where $\bar{x}$ is a fixed value, say m. Then, it is immediate to check that the extrema of the function $\log Z_v(\gamma, m)$ occur for γ^* such that $m = E_{v-1}[X]$:

Property 8. *If $\bar{x}$ is a fixed value m, then*

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\log Z_v(\gamma,m)\bigg|_{\gamma=\gamma^*} = 0. \tag{17}$$

if and only if γ^* is such that $m = E_{v-1}[X]$.

This result is important because it provides an easy way to find the value of the parameter γ of the optimum distributions (5) that solves the maximum entropy problems (4).

Proposition 9. *The values γ^* of the parameter γ of the optimum distributions that solve the maximum entropy problems (4) are the minimum of the maximizers of*

$$D_C(\gamma) = -\log Z_{\xi+1}(\gamma,m), \tag{18}$$

$$D_G(\gamma) = -\log Z_{-\xi}(\gamma,m), \tag{19}$$

where the two partitions functions involved are convex, possibly on several well defined intervals. Then, the entropy functionals $\mathscr{F}_\alpha^{(\cdot)}$ are simply given by

$$\mathscr{F}_\alpha^{(C \text{ resp. } G)}(m) = D_{C \text{ resp. } G}(\gamma^*). \tag{20}$$

Proof. Indeed, Theorem 1 and its corrolary indicates that the solution for the classical constraint (C) is obtained for $\bar{x} = m = E_\xi[X]$ and by $\bar{x} = m = E_{-\xi-1}[X]$ for the generalized constraint (G). Then by Property 8 it suffices to look for the extrema of $D_C(\gamma) = -\log Z_{\xi+1}(\gamma,m)$ in the first case or of $D_G(\gamma) = -\log Z_{-\xi}(\gamma,m)$ in the second case. With similar conditions of derivation as in Proposition 6 the second derivative of the partition function with respect to γ writes

$$\frac{\mathrm{d}^2 Z_v(\gamma,m)}{\mathrm{d}\gamma^2} = v(v-1)\int_{\mathscr{D}} (x-m)^2[1+\gamma(x-\bar{x}_\gamma)]^{v-2}Q(x)\,\mathrm{d}x, \tag{21}$$

$$= v(v-1)E_{v-2}[(X-m)^2]Z_{v-2}(\gamma,m). \tag{22}$$

For $v=\xi+1$ and $v=-\xi$, the factor $v(v-1)$ reduces to $\frac{\alpha}{(\alpha-1)^2}$. Since α is positive, the second derivative is always positive and the partition functions $Z_{\xi+1}(\gamma,m)$ and $Z_{-\xi}(\gamma,m)$ are convex on their domain of definition. On these domains, the functionals in (18) and (19) are then unimodal and their extrema are maxima.

In the discrete case and for $v<0$, $Z_v(\gamma,m)$ has singularities for all $\gamma=\frac{1}{m-k}$, where k is an integer in the support of the distribution. Therefore, $Z_v(\gamma,m)$ is only defined on segments $(\frac{1}{m-k},\frac{1}{m-k-1})$, for $m\notin(k+1,k))$, and $(\frac{1}{m-k-1},\frac{1}{m-k})$ for $m\in(k+1,k)$. In such a case, $-\log Z_v(\gamma,m)$ may present several maxima. The situation $v<0$ occurs for the classical constraint when $\alpha\in(0,1)$ (since the index $\xi+1=\alpha/(\alpha-1)$ is negative), and for the generalized constraint when $\alpha>1$. An example of functional $D_C(\gamma)$ with $\alpha=0.5$ in the case of a Poisson distribution is reported in Fig. 6. In the $v>0$ discrete case or in the continuous case, there is a single maximum.

Finally, since the expression of the Rényi information divergence of the optimum distributions is precisely the opposite of the log-partition function as indicated in Property 5, the value of functionals (18) and (19) at their optima γ^* such that $\bar{x}=m$ is precisely the value of entropy functionals $\mathscr{F}_\alpha^{(1)}(m)$ and $\mathscr{F}_\alpha^{(\alpha)}(m)$. □

Remark 10. When α tends to 1, the parameter $\tilde{\gamma}^*$ is thus the maximizer of (11), and we obtain

$$\lim_{\alpha\to 1}\mathscr{F}_\alpha^{(\cdot)} = \sup_{\tilde{\gamma}}\left\{\tilde{\gamma}\bar{x}-\log\int_{\mathscr{D}}\mathrm{e}^{\tilde{\gamma}x}Q(x)\,\mathrm{d}x\right\}, \tag{23}$$

that is the Cramér transform of $Q(x)$.

With the help of these different results it is now possible to characterize more precisely the entropy functionals.

Proposition 11. *Entropy functionals $\mathscr{F}_\alpha^{(C)}(m)$ and $\mathscr{F}_\alpha^{(G)}(m)$ are nonnegative, with an unique minimum at m_Q, the mean of Q, and $\mathscr{F}_\alpha^{(\cdot)}(m_Q)=0$. Furthermore, $\mathscr{F}_\alpha^{(C)}(m)$ is strictly convex for $\alpha\in[0,1]$.*

Proof. Rényi information divergence $D_\alpha(P||Q)$ is always nonnegative, and equal to zero for $P=Q$. Since functionals $\mathscr{F}_\alpha^{(\cdot)}(x)$ are defined as the minimum of $D_\alpha(P||Q)$, they are always nonnegative. If $P=Q$, we have also $P^*=Q$ and $m=E_P[X]=E_{P^*}[X]=m_Q$. Therefore $\mathscr{F}_\alpha^{(\cdot)}(m_Q)=0$ and m_Q is a global minimum.

From (16), we have $\frac{\mathrm{d}}{\mathrm{d}\bar{x}}\log Z_{v+1}(\gamma,\bar{x})=-\gamma(v+1)$. Then, functionals $\mathscr{F}_\alpha^{(\cdot)}(x)$ are only minimum if $\gamma=0$, and the corresponding optimum probability distributions are simply $P=Q$, and $D_\alpha(Q||Q)=0$. Therefore, $\mathscr{F}_\alpha^{(\cdot)}(x)$ have an unique minimum for $x=m_Q$, the mean of Q, and $\mathscr{F}_\alpha^{(\cdot)}(m_Q)=0$.

Finally, we examine the convexity of $\mathscr{F}_\alpha^{(C)}(m)$, for $\alpha\in[0,1]$.

Let P_1 and P_2 be the distributions that achieve the minimization of $D_\alpha(P||Q)$ subject to the constraints $x_1=E_P[X]$ and $x_2=E_P[X]$ respectively. Then, $\mathscr{F}_\alpha^{(C)}(x_1)=D_\alpha(P_1||Q)$, and $\mathscr{F}_\alpha^{(C)}(x_2)=D_\alpha(P_2||Q)$. In the same way, denote $\mathscr{F}_\alpha^{(C)}(\mu x_1+(1-\mu)x_2)=D_\alpha(\hat{P}||Q)$, where $\hat{P}$ denotes the optimum distribution with mean $\mu x_1+(1-\mu)x_2$. Distributions $\hat{P}(u)$ and $\mu P_1(u)+(1-\mu)P_2(u)$ have the same mean $\mu x_1+(1-\mu)x_2$. Hence, when $D_\alpha(P||Q)$ is a convex function of P, that is for $\alpha\in[0,1]$, we have $D_\alpha(P^*||Q)\leqslant\mu D_\alpha(P_1(u)||Q)+(1-\mu)D_\alpha(P_2(u)||Q)$, that is $\mathscr{F}_\alpha^{(C)}(\mu x_1+(1-\mu)x_2)\leqslant\mu\mathscr{F}_\alpha^{(C)}(x_1)+(1-\mu)\mathscr{F}_\alpha^{(C)}(x_2)$ and $\mathscr{F}_\alpha^{(C)}(x)$ is a convex function. □

Up to now the two optimization problems have been considered in parallel. But there is a special symmetry that enables to relate the solutions of the minimization of Rényi divergence subject to classical and generalized mean constraints. Then, there exists a simple relationship between the entropy functionals $\mathscr{F}_\alpha^{(C)}(x)$ and $\mathscr{F}_\alpha^{(G)}(x)$. Let us consider our original Rényi divergence minimization problem, on one hand with index α_1 and subject to a classical mean constraint m, and on the other hand with index α_2 and subject to a generalized

mean constraint m. The associated functionals, by Property 9, are $D_C(\gamma) = -\log Z_{\xi_1+1}(\gamma, m)$ and $D_G(\gamma) = -\log Z_{-\xi_2}(\gamma, m)$. Thus, we will have pointwise equality of these functions if $\xi_1 + 1 = -\xi_2$, that is if indexes α_1 and α_2 satisfy $\alpha_1 = 1/\alpha_2$. In this case, we will of course have equality of the optimum parameters γ, and the two optimization problems will have the same optimum value. Because of the pointwise equality functions $D_G(\gamma)$ and $D_G(\gamma)$, it is clear that the associated divergences are equal at the optimum, that is $D_{\alpha_1}(P_C||Q) = D_{\alpha_2}(P_G||Q)$. Besides this is easily checked in the general case: for the escort distribution $P^*(x)$ in (3), we *always* have the equality $D_{\frac{1}{\alpha}}(P^*||Q) = D_\alpha(P_1||Q)$. Hence, the minimization of the α Rényi divergence subject to the generalized mean constraint is exactly equivalent to the minimization of the $1/\alpha$ Rényi divergence subject to the classical mean constraint

$$\begin{cases} \inf\limits_{P_1} \; D_\alpha(P_1||Q) \\ \text{s.t. } E_{P^*}[X] = m \end{cases} = \begin{cases} \inf\limits_{P^*} D_{\frac{1}{\alpha}}(P^*||Q) \\ \text{s.t. } E_{P^*}[X] = m, \end{cases} \tag{24}$$

so that generalized and classical mean constraints can always be swapped, provided the index α is changed into $1/\alpha$, as was argued in [31,28]. Hence, equality (24) enables us to complete the characterization of entropy functionals $\mathscr{F}_\alpha^{(C)}(m)$ and $\mathscr{F}_\alpha^{(G)}(m)$:

Property 12. *Entropy functionals $\mathscr{F}_\alpha^{(C)}(m)$ and $\mathscr{F}_\alpha^{(G)}(m)$ admit the symmetry $\mathscr{F}_\alpha^{(G)}(x) = \mathscr{F}_{1/\alpha}^{(C)}(x)$. Besides, $\mathscr{F}_\alpha^{(C)}(m)$ is strictly convex for $\alpha \in [0, 1]$ and $\mathscr{F}_\alpha^{(G)}(m)$ is strictly convex for $\alpha \in [1, +\infty]$.*

Interestingly, it is also possible to define a *divergence in the object space*, that is a kind of generalized distance between two 'objects'. These divergences may be used for instance in clustering [30]. The objects are here considered as generalized means of distributions with minimum divergence to a reference measure $Q(x)$.

Proposition 13. *If P_1 and P_2 are two distributions in (5) with exponent $v = -\xi$ (generalized constraint), with $P_2 \ll P_1$, and with respective parameters γ_1, γ_2 and means m_1, m_2, then*

$$\mathscr{F}_\alpha^{(G)}(m_2, m_1) = D_\alpha(P_2||P_1) = \mathscr{F}_\alpha^{(G)}(m_2) - \mathscr{F}_\alpha^{(G)}(m_1) + \frac{1}{\alpha - 1} \log\left(1 - (\alpha - 1)\frac{\mathrm{d}\mathscr{F}_\alpha^{(G)}}{\mathrm{d}m}(m_1)(m_2 - m_1)\right), \tag{25}$$

and $\mathscr{F}_\alpha^{(G)}(m_2, m_1) \geqslant 0$, with equality if and only if $m_2 = m_1$.

Proof. The result is obtained by simple computations. First, we have

$$D_\alpha(P_2||P_1) = \frac{1}{\alpha - 1} \log \int \frac{[1 + \gamma_2(x - m_2)]^{\frac{\alpha}{1-\alpha}}}{Z_{-\xi}(\gamma_2, m_2)^\alpha} \frac{[1 + \gamma_1(x - m_1)]}{Z_{-\xi}(\gamma_1, m_1)^{1-\alpha}} Q(x)\,\mathrm{d}x$$

which can be rewritten as

$$D_\alpha(P_2||P_1) = \frac{1}{1 - \alpha}(\alpha \log Z_{-\xi}(\gamma_2, m_2) + (1 - \alpha)\log Z_{-\xi}(\gamma_1, m_1) - \log Z_{-\xi-1}(\gamma_2, m_2)) \tag{26}$$

$$+ \frac{1}{\alpha - 1} \log\left[1 + \gamma_1 \int (x - m_1) \frac{[1 + \gamma_2(x - m_2)]^{-\xi-1}}{Z_{-\xi-1}(\gamma_2, m_2)} Q(x)\,\mathrm{d}x\right]. \tag{27}$$

In the first line, we have $Z_{-(\xi+1)}(\gamma_2, m_2) = Z_{-\xi}(\gamma_2, m_2)$ by Property 4, Eq. (13), and we recognize from Proposition 9 that $\mathscr{F}_\alpha^{(G)}(m) = -\log Z_{-\xi}(\gamma, m)$. In the second line, the integral reduces to $(m_2 - m_1)$ since m_2 is the generalized mean of the distribution P_2. Finally, γ_1 can be expressed as the derivative of the log-partition function as stated by (16) in Property 7.

By definition, $\mathscr{F}_\alpha^{(G)}(m_2, m_1)$ is the Rényi information divergence $D_\alpha(P_2||P_1)$ which is always greater or equal to zero, with equality if and only if $P_2 = P_1$, which implies $m_2 = m_1$. □

For $\alpha \to 1$, $\mathscr{F}_\alpha^{(G)}(m_2, m_1)$ reduces to a standard *Bregman divergence*. Indeed, using $\log(1 - x) \simeq -x$, we have simply

$$\lim_{\alpha \to 1} \mathscr{F}_\alpha^{(G)}(m_2, m_1) = \mathscr{F}_\alpha^{(G)}(m_2) - \mathscr{F}_\alpha^{(G)}(m_1) - \frac{\mathrm{d}\mathscr{F}_\alpha^{(G)}}{\mathrm{d}m}(m_1)(m_2 - m_1).$$

3. Examples of entropy functionals

We now examine four special cases for the reference measure $Q(x)$: a uniform and an exponential distribution that model systems with continuous states; and then a Bernoulli (two-levels) and a Poisson distribution which may model systems with discrete states. The minima of the Rényi divergence, that is the entropies $\mathscr{F}_\alpha^{(C \text{ or } G)}(x)$, are attained for the values γ^* that maximize the functionals $D_C(\gamma)$ and $D_G(\gamma)$ in Proposition 9. This involves the computation of $Z_v(\gamma, m)$ for all reference measures Q considered, and the resolution of $\frac{\mathrm{d}}{\mathrm{d}\gamma} Z_{v+1}(\gamma, m) = 0$. The case $\alpha = 1$ is obtained in the limit $|\, v \,| \to +\infty$, since $|\, \xi \,| \to +\infty$ when α tends to 1. Results of numerical evaluations for varying α are provided.

3.1. Uniform reference

Let us first consider the case of the uniform reference $Q(x)$ on $[0,1]$. The partition function is given by $Z_v(\gamma, m) = \int_{\mathscr{D}} [\gamma(x-m)+1]^v \,\mathrm{d}x$, where the domain $\mathscr{D}$ is defined by $\mathscr{D} = \mathscr{D}_Q \cap \mathscr{D}_\gamma$, with $\mathscr{D}_Q = \{x : x \in [0,1]\}$ and $\mathscr{D}_\gamma = \{x : \gamma(x-m)+1 \geqslant 0\}$.

Computation of the partition function in the different domains together with the fact that $m \in [0,1]$ leads to

$$Z_v(\gamma, m) = \left(\frac{1}{\gamma(1+v)} (\gamma - \gamma m + 1)^{v+1} U\left(\gamma - \frac{1}{m-1}\right) - (-\gamma m + 1)^{v+1} U\left(-\gamma + \frac{1}{m}\right) \right)$$

$$\text{for all } \gamma \text{ if } v \geqslant 0, \text{ for } \gamma \in \left(\frac{1}{m-1}, \frac{1}{m}\right) \text{ if } v < 0, \text{ and } Z_v(\gamma, m) = +\infty \text{ otherwise,}$$

where U denotes the Heaviside distribution: $U(t) = 0$ for $t < 0$ and $U(t) = 1$ for $t > 0$.

The first derivative of the partition function is given by

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} Z_v(\gamma, m) = -\frac{v\gamma(m-1)+1}{\gamma^2(v+1)} (\gamma(m-1)+1)^v U\left(\gamma - \frac{1}{m-1}\right) + \frac{\gamma m(v)+1}{\gamma^2(v+1)} (1-\gamma m)^v U\left(-\gamma + \frac{1}{m}\right). \tag{28}$$

We next have to look for the expression of entropy functionals $\mathscr{F}_\alpha^{(\cdot)}(x)$. Unfortunately, no analytical solution can be exhibited here, but the two functionals still can be evaluated numerically. For the classical mean constraint (C) we can check that $\mathscr{F}_\alpha^{(C)}(x)$ is a family of convex functions on $(0,1)$, minimum for the mean of the reference measure Q, as was indicated in Proposition 11. In the same way, we can check that for the generalized mean constraint (G) $\mathscr{F}_\alpha^{(G)}(x)$ is a family of nonnegative functions on $(0,1)$, also minimum for the mean of

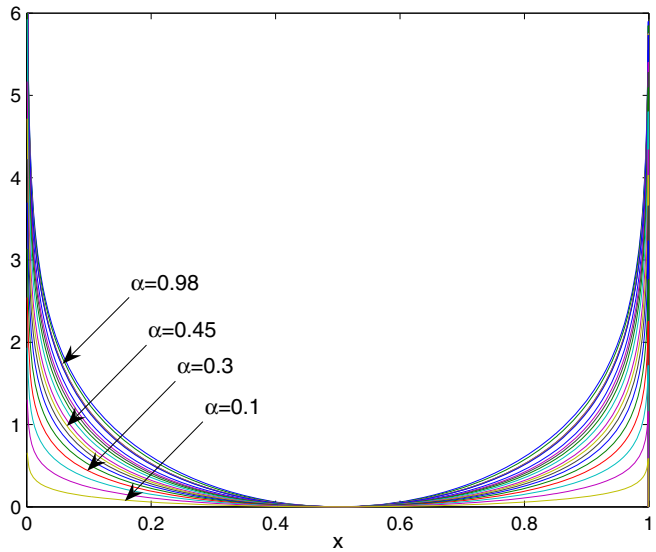


Fig. 1. Entropy functional $\mathscr{F}_\alpha^{(C)}(x)$ for a uniform reference measure and $\alpha \in (0,1)$.

the reference measure Q. The entropies $\mathscr{F}_\alpha^{(C)}(x)$ and $\mathscr{F}_\alpha^{(G)}(x)$ were evaluated numerically and are given in Figs. 1 and 2 for $\alpha \in (0,1)$. Of course, the $\alpha \leftrightarrow 1/\alpha$ duality given in Property 12 enables to extend these two functionals for $\alpha > 1$.

Hence, it is apparent that the minimization of $\mathscr{F}_\alpha^{(\cdot)}(x)$ under some constraint would automatically lead to a solution on $(0,1)$. Moreover, the parameter α may serve to tune the curvature of the functional and the degree of penalization of bounds.

3.2. Exponential reference

The exponential probability density function is $Q(x) = \beta e^{-\beta x}$, for $x \geqslant 0$ and $\beta > 0$. The partition function is given by

$$Z_v(\gamma, m) = \beta \int_{\mathscr{D}} [\gamma(x-m)+1]^v e^{-\beta x}\, dx, \tag{29}$$

where $\mathscr{D} = \{x : x \geqslant \max\{0, m - \frac{1}{\gamma}\}$ if $\gamma > 0$ or $x \in [0, m - \frac{1}{\gamma}]$ if $\gamma < 0\}$, ensures that the integrand $[\gamma(x-m)+1]$ is nonnegative and the integral finite.

The evaluation of $Z_v(\gamma, m)$ on the different domains gives

$$Z_v(\gamma, m) = \begin{cases} e^{-\beta\frac{\gamma m-1}{\gamma}}\left(\frac{\gamma}{\beta}\right)^v \Gamma(v+1) & \text{if } \gamma > \frac{1}{m} > 0, \\ & \quad v \geqslant 0, \\ e^{-\beta\frac{\gamma m-1}{\gamma}}\left(\frac{\gamma}{\beta}\right)^v \Gamma(v+1, \beta\frac{1-\gamma m}{\gamma}) & \text{if } \frac{1}{m} > \gamma > 0, \\ e^{-\beta\frac{\gamma m-1}{\gamma}}\left(\frac{\beta}{\gamma}\right)^{-v}\left(\Gamma\left(v+1, \beta\frac{-\gamma m+1}{\gamma}\right) - \Gamma(v+1)\right) & \text{if } \gamma < 0 < \frac{1}{m}, \\ & \quad v \geqslant 0, \end{cases} \tag{30}$$

and $Z_v(\gamma, m) = +\infty$ for $\gamma < 0$ or $\gamma > \frac{1}{m}$ if $v < 0$. Let us now examine the behavior of the entropies $\mathscr{F}_\alpha^{(\cdot)}(x)$ when $\alpha \to 1$. This amounts to study $Z_v(\gamma, m)$ and its maximum when $|\, v \,| \to +\infty$. The simplest derivation is as follows. As in Remark 3, let $\gamma = \tilde{\gamma}/v$, so that $(1+\gamma(x-m))^v \sim \exp(\tilde{\gamma}(x-m))$. In this case, one easily obtain that

$$\log Z_v(\tilde{\gamma}, m) \simeq \log\beta - \tilde{\gamma}m - \log(\beta - \tilde{\gamma}), \tag{31}$$

whose derivative is equal to zero for

$$\tilde{\gamma}^* = \beta - \frac{1}{m}. \tag{32}$$

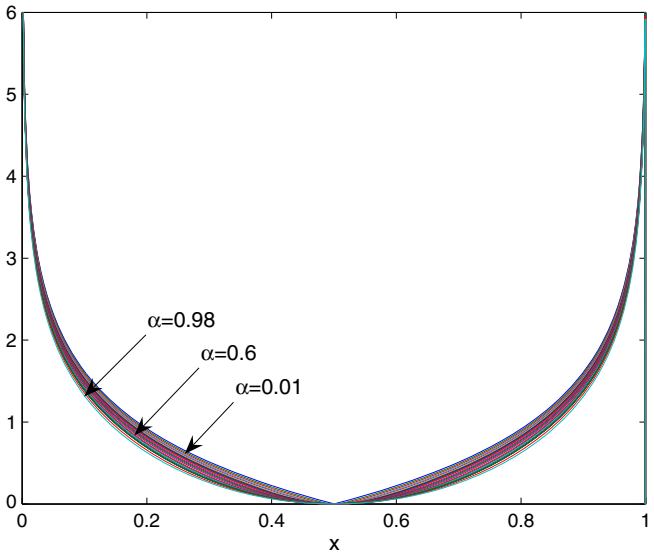


Fig. 2. Entropy functional $\mathscr{F}_\alpha^{(G)}(x)$ for a uniform reference measure and $\alpha \in (0,1)$.

We shall also note that if $v < 0$, the sign of $\gamma = \tilde{\gamma}/v$ is the sign of $(1 - \beta m)$. Since $Z_v(\gamma, m)$ is only defined for $\gamma > 0$ when $v < 0$, it means that we only have a solution for $m < 1/\beta$. Indeed, for $\gamma > 0$ and $v < 0$, the factor $(1 + \gamma(x - m))^v$ is decreasing, and consequently the mean of the optimum distribution (5) cannot be greater than the mean of the reference distribution, $m_Q = E_Q[X] = 1/\beta$.

With the optimum value $\tilde{\gamma}^*$, the log partition function becomes

$$\log Z_v(\gamma^*, m) \simeq -(\beta m - 1) + \log(\beta m) \quad (\forall\ m \text{ if } v \to +\infty, \text{ for } m < 1/\beta \text{ if } v \to -\infty). \tag{33}$$

Finally, we thus obtain

$$\mathscr{F}^{(C)}_{\alpha\to 1}(x) = -\log Z_{\xi+1}(\gamma^*, x) = (\beta x - 1) - \log(\beta x) \tag{34}$$

for $x < 1/\beta$ when α tends to 1 by lower values, and for all x if α tends to 1 by higher values. By the duality Property 12, this expression is also the limit form of functional $\mathscr{F}^{(G)}_{\alpha}(x)$. As was expected, the functional $(\beta x - 1) - \log(\beta x)$ is strictly convex, positive and zero for $x = 1/\beta$, the mean of the exponential distribution. It was employed in speech processing and is called the *Itakura-Saïto entropy functional.* For $\beta = 1$, it reduces to the so-called *Burg entropy* that is well-known in spectrum analysis. The entropy functionals can be evaluated numerically. For instance, $\mathscr{F}^{(G)}_{\alpha}(x)$ is given on Fig. 3 for $\alpha > 0$. It is a family of nonnegative functions, equal to zero for $x = m_Q = 1/\beta$, and convex for $\alpha \in [1, +\infty)$.

3.3. Bernoulli reference

Let us now consider the case of the Bernoulli measure $Q(x) = \beta\delta(x) + (1 - \beta)\delta(x - 1)$. Of course, the (generalized) mean of optimum distributions is somewhere in the interval $[0, 1]$. When γ is outside of the interval $(\frac{1}{m-1}, \frac{1}{m})$, the probability distribution reduces to a pure state – $\delta(x)$ or $\delta(x - 1)$, and its (generalized) mean is 0 or 1. Incorporation of the bounds into the domain depends on the sign of v : for $v < 0$, $Z_v(\gamma, m)$ diverges to $+\infty$ on the bounds whereas it remains finite for $v > 0$. The expression of the partition function follows directly from the definition:

$$Z_v(\gamma, m) = \beta(1 - \gamma m)^v + (1 - \beta)(1 + \gamma(1 - m))^v. \tag{35}$$

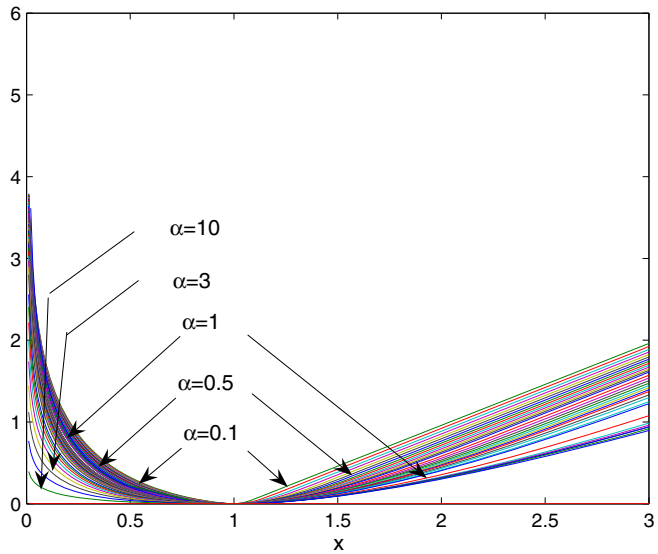


Fig. 3. Entropy functional $\mathscr{F}^{(G)}_{\alpha}(x)$ for an exponential reference measure with $\beta = 1$ and $\alpha > 0$. By Property 12 it is also $\mathscr{F}^{(C)}_{1/\alpha}(x)$.

In contrast to the previous case, it is possible here to obtain an explicit expression of the entropy functionals for any α. Indeed, if p denotes the value of the optimum distribution at $x = 1$, then the generalized expectation is

$$m = \frac{\sum_{x=0}^{1} xP(x)^{\alpha}Q(x)^{1-\alpha}}{\sum_{x=0}^{1} P(x)^{\alpha}Q(x)^{1-\alpha}} = \frac{(1-\beta)^{1-\alpha}p^{\alpha}}{\beta^{1-\alpha}(1-p)^{\alpha} + (1-\beta)^{1-\alpha}p^{\alpha}} \tag{36}$$

and it is therefore possible to express p as a function of m

$$p = \frac{\left(\beta^{1-\alpha}m\right)^{\frac{1}{\alpha}}}{\left(\beta^{1-\alpha}m\right)^{\frac{1}{\alpha}} + \left((1-\beta)^{1-\alpha}(1-m)\right)^{\frac{1}{\alpha}}}. \tag{37}$$

Now, since the Rényi information divergence is

$$D_{\alpha}(P||Q) = \frac{1}{\alpha - 1}\log[\beta^{1-\alpha}(1-p)^{\alpha} + (1-\beta)^{1-\alpha}p^{\alpha}] \tag{38}$$

it suffices to replace p by the expression (37) which leads to

$$\mathscr{F}_{\alpha}^{(G)}(m) = \frac{\alpha}{1-\alpha}\log\left[\beta^{1-\frac{1}{\alpha}}(1-m)^{\frac{1}{\alpha}} + (1-\beta)^{1-\frac{1}{\alpha}}m^{\frac{1}{\alpha}}\right]. \tag{39}$$

The case of the classical mean is even simpler: we have $m = p$, and $\mathscr{F}_{\alpha}^{(C)}(m)$ has the expression of the divergence in (38) with p replaced by m. It is also interesting to note, and check, that the $\alpha \leftrightarrow 1/\alpha$ duality of Property 12 links these two expressions.

The limit case $\alpha \to 1$ is easily derived using L'Hospital's rule. It comes

$$\mathscr{F}_{\alpha\to 1}^{(\cdot)}(x) = x\ln\left(\frac{x}{1-\beta}\right) + (1-x)\ln\left(\frac{1-x}{\beta}\right). \tag{40}$$

This expression is the celebrated *Fermi–Dirac entropy* that is strictly convex, nonnegative, and equal to zero for $x = E_Q[X] = 1 - \beta$, the mean m_Q of the reference measure.

Plots of the entropy functionals are given in Figs. 4 and 5 for $\alpha \in (0, 1)$ and $\beta = 1/2$. In both cases, we have a family of nonnegative functions, equal to zero for the mean of the reference measure. It can also be checked that $\mathscr{F}_{\alpha}^{(C)}(x)$ is convex for $\alpha \in (0, 1]$.

3.4. Poisson reference

As a final example, let us consider the case of a Poisson measure $Q(x) = \frac{\mu^x}{x!}\mathrm{e}^{-\mu}$, for $x \geqslant 0$. Domain $\mathscr{D}$ is $\mathscr{D} = \mathscr{D}_Q \cap \mathscr{D}_\gamma$, where $\mathscr{D}_Q = \mathbb{N}^+$ and $\mathscr{D}_\gamma = \{x : \gamma(x-m)+1 \geqslant 0\}$. The partition function is given by

$$Z_v(\gamma, m) = \sum_{\mathscr{D}} [\gamma(x-m)+1]^v \frac{\mu^x}{x!}\mathrm{e}^{-\mu}. \tag{41}$$

Three cases appear, according to the value of γ:

(a) if $\frac{1}{m} \geqslant \gamma \geqslant 0$, then $\mathscr{D}$ reduces to $\mathscr{D}_1 = \{x : x \in [0, +\infty)\}$;
(b) for $\gamma \geqslant \frac{1}{m}$ the domain is $\mathscr{D}_2 = \{x : x \in [\lceil m - \frac{1}{\gamma}\rceil, +\infty)\}$;
(c) when $\gamma < 0$, $\mathscr{D} = \mathscr{D}_3 = \{x \in [0, \lfloor m - \frac{1}{\gamma}\rfloor]\}$.

In these expressions $\lfloor x\rfloor$ denotes the floor function that returns the largest integer less than or equal to x; and $\lceil x\rceil$ is the ceil function, the smallest integer not less than x. Closed-form formulas can not be derived in the general case, but only in the case of an integer exponent v. When v is not an integer, we will have to resort to the serie (41), possibly truncated for numerical computations. In order to save space, we only sketch the derivation in $\mathscr{D}_1$:

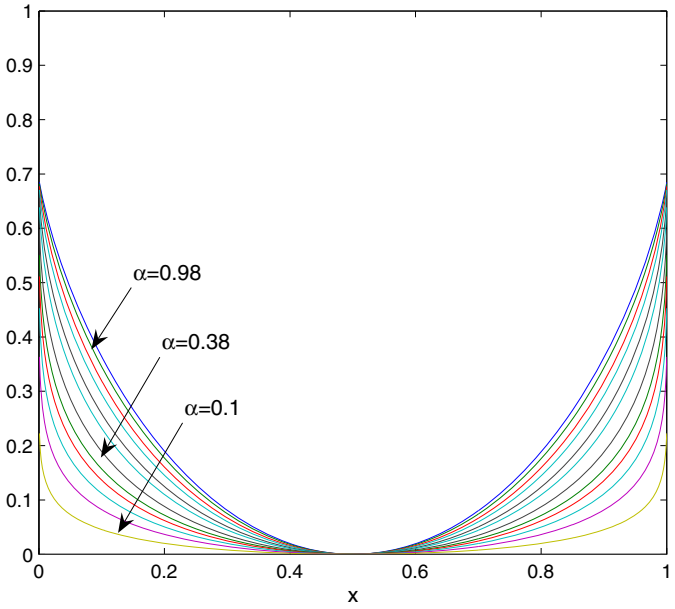


Fig. 4. Entropy functional $\mathscr{F}_{\alpha}^{(C)}(x)$ for a Bernoulli reference measure and $\alpha \in (0,1)$.

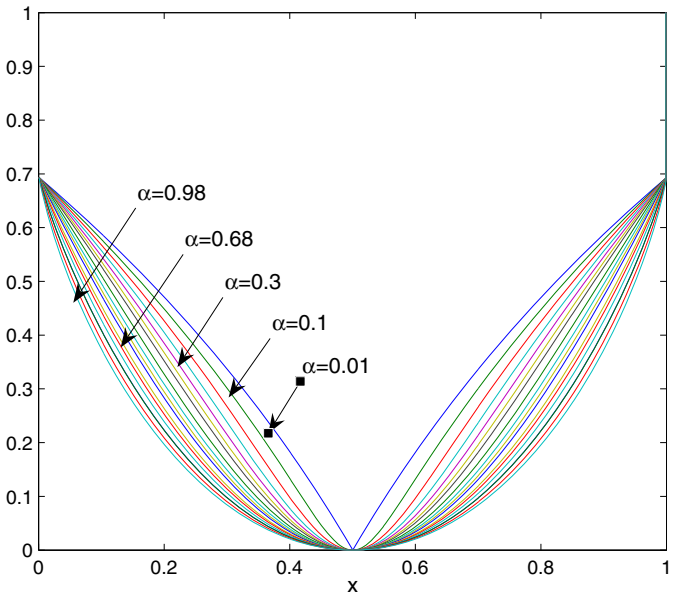


Fig. 5. Entropy functional $\mathscr{F}_{\alpha}^{(G)}(x)$ for a Bernoulli reference measure and $\alpha \in (0,1)$.

$$Z_v(\gamma, m) = (1-\gamma m)^v \mathrm{e}^{-\mu} \sum_{x=0}^{+\infty} (\theta x + 1)^v \frac{\mu^x}{x!} \tag{42}$$

with $\theta = \frac{\gamma}{1-\gamma m}$. In the serie above the ratio of successive terms $\frac{(1+\theta x+\theta))^v}{(x+1)(1+\theta x)^v}\mu$ is the ratio of two completely factored polynomials. This indicates that the serie can be written as a generalized hypergeometric function, when v is integer. So doing, we obtain

$$Z_v(\gamma, m) = (1-\gamma m)^v \mathrm{e}^{-\mu}_{|v|} F_{|v|}(a, \ldots, a, b, \ldots, b, \mu)$$

with $a = (1+\theta)/\theta$ and $b = 1/\theta$ for $v > 0$; or with $a = 1/\theta$ and $b = (1+\theta)/\theta$ for $v < 0$.

The derivative with respect to γ is

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} Z_{v+1}(\gamma, m) = (1-\gamma m)^v \mathrm{e}^{-\mu}(v+1) \sum_{x=0}^{+\infty} (x-m)(1+\theta x)^v \frac{\mu^x}{x!}, \tag{43}$$

that can also be expressed using hypergeometric functions. Formulas for domains $\mathscr{D}_2$ and $\mathscr{D}_3$ also involve hypergeometric functions. With these formulas, or by direct evaluation of (41), functionals $D_C(\gamma)$ and $D_G(\gamma)$ can be evaluated and maximized on their domains of definition so as to find the optimum value γ^*.

Given the signs of v and γ, and the supports $\mathscr{D}_1$, $\mathscr{D}_2$ and $\mathscr{D}_3$, it is already possible to deduce that the solution γ^* is necessarily in a specific interval. Hence, we obtain here that for $v > 0$ (respectively for $v < 0$), solutions associated to a constraint $m > \mu$ corresponds to case (a) (resp. case (c)) and that solutions for $m < \mu$ correspond to case (c) (resp. case (a)). The argumentation relies on the fact that if P_i and P_j are two optimum distributions with supports $\mathscr{D}_i$ and $\mathscr{D}_j$, with the same (generalized) mean but different parameters, then by Theorem 1 $D_\alpha(P_j||Q) \geqslant D_\alpha(P_i||Q)$ if P_j is dominated by P_i.

In the case $m > \mu$, $v < 0$, the solution with minimum divergence is for a distribution P_3 in case (c), and furthermore we have $D_\alpha(P_3||Q) \to 0$. This can be seen as follows. Let $x \in \mathscr{D}_3$ and $k = \lfloor m - \frac{1}{\gamma} \rfloor$, so that $x < k+1$. Let now $\gamma = \frac{1}{m-k} + \epsilon$ with $\epsilon \in (0, \frac{1}{m-k-1} - \frac{1}{m-k})$. Then the mean of the distribution is given by

$$E_v[X] = \frac{1}{Z_v(\gamma, m)} \sum_{x=0}^{k-1} x \left[\frac{k-x}{k-m}\right]^v Q(x) + k \frac{[k-m]^v}{Z_v(\gamma, m)} \epsilon^v Q(k), \tag{44}$$

and any value higher than $\mu = E_Q[X]$ can be obtained by tuning ϵ, for many values of k. When k increases, $\gamma = \frac{1}{m-k}$ tends to 0 by lower values and P_3 tends to Q, which results in $D_\alpha(P_3||Q) \to 0$.

The $v < 0$ case has the specificity that $Z_v(\gamma, m)$ exhibits singularities at $\gamma = \frac{1}{m-k}$ for all $k \geqslant 0$. Then $Z_v(\gamma, m)$, with $v = -(\xi+1)$ or $v = \xi$, is only convex on intervals $[\frac{1}{m-k}, \frac{1}{m-k-1}]$ or $[\frac{1}{m-k-1}, \frac{1}{m-k}]$ (for $k+1 > m > k$), with $Z_v(\gamma, m) = +\infty$ on the bounds of each interval. Consequently, $-\log Z_v(\gamma, m)$ may present several maxima. This is illustrated in Fig. 6 where function $D_C(\gamma)$ with $\alpha = 0.5$ presents many extrema. The solution with minimum Rényi divergence corresponds to the minimum of these maxima. The limit case $\alpha \to 1$ is obtained with $|v| = |\xi| \to +\infty$. According to the discussion above, the optimum γ corresponds to case (a) for $\{m > \mu, v > 0\}$ and $\{m < \mu, v < 0\}$, and to case (c) for $\{m < \mu, v > 0\}$. For case (a), the support is $\mathscr{D}_1$, and the derivative of the partition function $Z_v(\gamma, m)$ is given by (43). In this derivative, the sum can be rewritten as

$$\sum_{x=0}^{+\infty} (x-m)(1+\theta x)^v \frac{\mu^x}{x!} = \sum_{x=0}^{+\infty} (\mu(1+\theta x+\theta)^v - m(1+\theta x)^v) \frac{\mu^x}{x!}, \tag{45}$$

so that $Z_{v+1}(\gamma, m)$ is minimum when the RHS of (45) is equal to zero. We have to solve this equation in θ. Suppose that θ is small and that $\theta x \ll 1$ for the significative values of the probability distribution. In this case, we use the approximation $(1+\theta x)^v = \mathrm{e}^{v \log(1+\theta x)} \approx \mathrm{e}^{v\theta x}$, that leads to

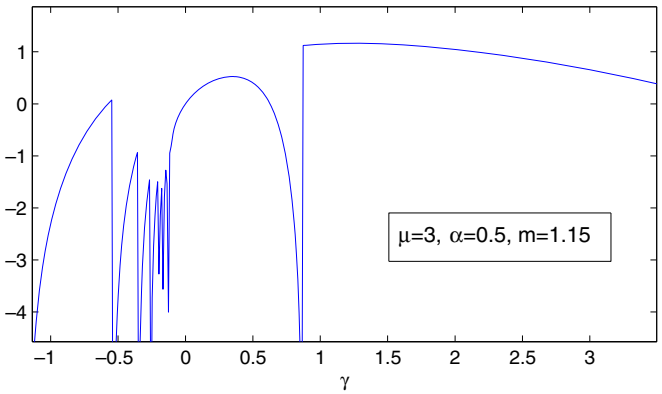


Fig. 6. Example of functional $D_C(\gamma)$ for the Poisson reference with classical mean constraint, with $\mu = 3, \alpha = 0.5$ ($\xi = -2$) and $m = 1.15$. It presents singularities at $1/(m-k), \forall k$, and maxima at $\gamma = 0.35$ and $\gamma = 1.24$.

$$\sum_{x=0}^{+\infty}(\mu e^{v\theta(x+1)} - m e^{v\theta x})\frac{\mu^x}{x!} = e^{\mu e^{v\theta}}(\mu e^{v\theta} - m) = 0. \tag{46}$$

The solution is given by $\theta^* = \frac{1}{v}\log(\frac{m}{\mu})$, that in turn provides

$$\gamma^* = \frac{\ln\frac{m}{\mu}}{v + m\ln\frac{m}{\mu}}. \tag{47}$$

In case (a), γ is positive, and this will be true for γ^* if $\{m > \mu, v > 0\}$ or $\{m < \mu, v < 0\}$. For the log-partition function, when $|v| \to +\infty$, this leads to

$$-\log\ Z_{v+1}(\gamma^*, m) \approx m\log\frac{m}{\mu} + (\mu - m). \tag{48}$$

In domain $\mathscr{D}_3$, the derivative of the partition function $Z_v(\gamma, m)$ is equal to zero if

$$\sum_{x=0}^{k}(x-m)(1+\theta x)^v\frac{\mu^x}{x!} = 0, \quad \text{with } k = \left\lfloor m - \frac{1}{\gamma}\right\rfloor,\ \gamma < 0.$$

If γ is small enough, $k \to +\infty$ and we obtain for $v > 0$ the same formulation and solution as in $\mathscr{D}_1$ The solution γ^* in (47) is now negative, that imposes $m < \mu$ for $v > 0$. Finally, we have shown above that if $m > \mu$ with $v < 0$ then $D_\alpha(P_3||Q) \to 0$.

Hence, we obtain that the entropy functionals converge to

$$\mathscr{F}^{(\cdot)}_{\alpha\to 1}(x) = x\ln\frac{x}{\mu} + (\mu - x) \tag{49}$$

with the restriction that $\mathscr{F}^{(\cdot)}_\alpha(x) = 0$ for $x > \mu$ if (C) $\alpha < 1$ or (G) $\alpha > 1$.

This functional is simply the cross-entropy between x and μ or Kullback–Leibler (Shannon) entropy functional with respect to μ [9]. It measures a 'distance' between a possible mean (observable) and a reference mean μ, and it has been used as a regularization functional in several applied problems, such as astronomy, tomography, RMN, and spectrometry.

As in the previous cases, the entropy functionals $\mathscr{F}^{(C)}_\alpha(x)$ and $\mathscr{F}^{(G)}_\alpha(x)$ can be evaluated numerically. For instance, $\mathscr{F}^{(G)}_\alpha(x)$ is given in Fig. 7 for $\mu = 3$. It presents an unique minimum for $m = \mu$, and we note that it is is not convex for small values of α.

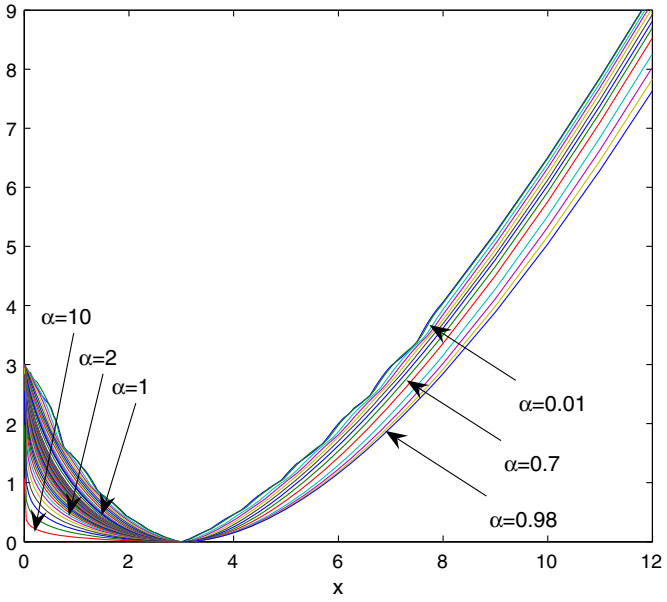


Fig. 7. Entropy functional $\mathscr{F}^{(G)}_\alpha(x)$ for a Poisson reference measure with $\mu = 3$ and $\alpha \geqslant 0$. For $\alpha \to 1$, $\mathscr{F}^{(\cdot)}_\alpha(x)$ converges to $x\ln\frac{x}{\mu} + (\mu - x)$.

4. Conclusion and future work

By weakening one of the postulates that lead to the definition of Shannon entropy, Rényi [32] introduced a one parameter family of entropy and divergence. Shannon entropy and Kullback–Leibler divergence are recovered in the limiting case for the parameter $\alpha \to 1$. In this work, we considered the maximum entropy problems associated with Rényi Q-entropies. We characterized the solutions for a standard mean constraint and for the generalized mean constraint of nonextensive statistics. We defined and discussed the entropy functionals as a function of the constraints. These entropies were characterized and various properties and relationships were highlighted. We also discussed numerical aspects. Finally we illustrated this setting through some specific examples and recovered some well-known entropy functionals.

Future work will consider the extension of this setting in the multivariate case. An issue that should be examined is the fact that the direct multivariate extension of (5) is not separable in the case of a separable reference $Q(x)$; which means that some dependences are implicitly introduced in the maximum entropy solution.

We also intend to investigate a possible underlying geometrical structure of the maximum entropy distributions (5). This structure should extend the geometrical structure of exponential families and involve the Bregman-like divergence introduced by (25).

Finally, maximum entropy methods have been successfully employed for solving inverse problems. We intend to consider the potential of Rényi entropies and divergence in this field. A simple contribution would be to examine the interest of a Rényi entropy functional, e.g. (39), as a potential in a Markov field for image deconvolution or restoration.

Appendix A

Proof of Theorem 1. Let us begin with the classical constraint (C). In this first case, we follow the approach of [37]. Consider the functional Bregman divergence:

$$B_h(f,g) = \int d(f,g)h(x)\,\mathrm{d}x = \int -(f(x)^\alpha - g(x)^\alpha - \alpha(f(x)-g(x))g(x)^{\alpha-1})h(x)\,\mathrm{d}x,$$

where $h(x)$ is a nonnegative functional, associated to the (pointwise) Bregman divergence $d(f,g)$ built upon the strictly convex function $-x^\alpha$ for $\alpha \in (0,1)$. Then

$$B_{Q^{1-\alpha}}(P,P_C) = -\int_{\mathscr{S}} P(x)^\alpha - P_C(x)^\alpha - \alpha(P(x)P_C(x)^{\alpha-1} - P_C(x)^\alpha)Q(x)^{1-\alpha}\,\mathrm{d}x, \tag{A.1}$$

$$= -\int_{\mathscr{S}} P(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x + \int_{\mathscr{S}} P_C(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x \tag{A.2}$$

with $h(x) = Q(x)^{1-\alpha}$ and where $\mathscr{S}$ denotes the support of $P_C(x)$. The second line follows from the fact that when P and P_C have the same mean $\bar{x} = E_{P_C}[X] = E_P[X]$, then using the expression in (5) with $v = \xi = \frac{1}{\alpha-1}$ it is possible to check that

$$\int_{\mathscr{S}} P(x)P_C(x)^{\alpha-1}Q(x)^{1-\alpha}\,\mathrm{d}x = \int_{\mathscr{S}} P_C(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x = Z_\xi(\gamma,\bar{x})^{-\alpha}$$

provided the whole support of $P(x)$ is included in $\mathscr{S}$, which is the case by the absolute continuity of $P(x)$ with respect to $P_C(x)$.

The Bregman divergence $B_{Q^{1-\alpha}}(P,P_C)$ being always positive and equal to zero if and only if $P = P_C$, the equality (A.2) implies that, for $\alpha \in (0,1)$,

$$D_\alpha(P||Q) \geqslant D_\alpha(P_C||Q) \tag{A.3}$$

which means that P_C is the distribution with minimum Rényi (Tsallis) divergence to Q, in the set of all distributions $P \ll P_C$ with a given mean $\bar{x}$, for $\alpha \in (0,1)$. The case $\alpha > 1$ can be derived accordingly, beginning with the Bregman divergence associated to the strictly convex function x^α.

As far as the generalized mean constraint (G) is concerned, let us now consider the Rényi information divergence $D_\alpha(P||P_G)$ from P to P_G, with P_G given in (5) with $v = -\xi = \frac{1}{1-\alpha}$

$$(\alpha-1)D_\alpha(P||P_G)=\log\int_{\mathscr{S}}P(x)^\alpha P_G(x)^{1-\alpha}\,\mathrm{d}x, \tag{A.4}$$

with $\mathscr{S}$ the support of $P_G(x)$, and which can be rearranged as

$$(\alpha-1)D_\alpha(P||P_G)=\log\int_{\mathscr{S}}\frac{P(x)^\alpha Q(x)^{1-\alpha}}{\int_{\mathscr{S}}P(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x}[\gamma(x-\bar{x})+1]\,\mathrm{d}x \tag{A.5}$$

$$+\log\int_{\mathscr{S}}P(x)^\alpha Q(x)^{1-\alpha}\,\mathrm{d}x-(1-\alpha)\log Z_{\frac{1}{1-\alpha}}(\gamma,\bar{x}). \tag{A.6}$$

The generalized mean with respect to P appears in the first term, and cancels if P and P_G have the same generalized mean $\bar{x}$ and $P_G\gg P$. In such a case, we obtain

$$D_\alpha(P||P_G)=\frac{1}{(\alpha-1)}\log\int P(x)^\alpha Q^{1-\alpha}\,\mathrm{d}x+\log Z_{\frac{1}{1-\alpha}}(\gamma,\bar{x}), \tag{A.7}$$

$$=D_\alpha(P||Q)-D_\alpha(P_G||Q), \tag{A.8}$$

where we used the fact that $D_\alpha(P_G||Q)=-\log Z_{\frac{1}{1-\alpha}}(\gamma,\bar{x})$ as stated in Proposition 5. Since the Rényi information divergence is always greater or equal to zero, we have

$$D_\alpha(P||Q)\geqslant D_\alpha(P_G||Q) \tag{A.9}$$

and conclude that P_G is the distribution with minimum Rényi (Tsallis) divergence to Q, in the set of all distributions $P\ll P_G$ with a given generalized α-mean $\bar{x}$. Finally, it is easy to check, given the expression of P_G and the fact that $\alpha\xi=\xi+1$, that the generalized mean of P_G is also the standard mean of the distribution with exponent $v=-(\xi+1)$, that is $E^{(\alpha)}_{P_G}[X]=E_{P^*_G}[X]=E_{-(\xi+1)}[X]$.

Note that the equality in (A.8), $D_\alpha(P||Q)=D_\alpha(P||P_G)+D_\alpha(P_G||Q)$, is a pythagorean equality, which means that P_G is the orthogonal projection of P on the set of probability distributions with fixed generalized mean $\bar{x}$.

Appendix B

Proof of Proposition 6. The exact behaviour depends on the reference distribution $Q(x)$ and on the sign of the exponent v. Because the domain of definition $\mathscr{D}$ might depend on γ, the derivative of the partition function writes

$$\frac{\mathrm{d}Z_v(\gamma,\bar{x}_\gamma)}{\mathrm{d}\gamma}=\lim_{\delta\gamma\to0}\frac{1}{\delta\gamma}(Z_v(\gamma+\delta\gamma,\bar{x}_{\gamma+\delta\gamma})-Z_v(\gamma,\bar{x}_\gamma)),$$

where $\bar{x}_\gamma$ and $\bar{x}_{\gamma+\delta\gamma}$ now denote the parameter $\bar{x}$ for distributions with parameter γ and $\gamma+\delta\gamma$. Let us begin with the continuous case. If $\delta\mathscr{D}$ denotes the domain increment associated to the variation $\delta\gamma$, it remains

$$\frac{\mathrm{d}Z_v(\gamma,\bar{x}_\gamma)}{\mathrm{d}\gamma}=\int_{\mathscr{D}}\frac{\mathrm{d}}{\mathrm{d}\gamma}(1+\gamma(x-\bar{x}_\gamma))^v Q(x)\,\mathrm{d}x \tag{B.1}$$

$$+\lim_{\delta\gamma\to0}\frac{1}{\delta\gamma}\int_{\delta\mathscr{D}}(1+(\gamma+\delta\gamma)(x-\bar{x}_{\gamma+\delta\gamma}))^v Q(x)\,\mathrm{d}x \tag{B.2}$$

Of course, when $\mathscr{D}$ does not depend on γ, we only have the first term, and it is easy to obtain (14). Otherwise, in order to satisfy the positivity of the integrand, the domain $\mathscr{D}$ is bounded above by $(\bar{x}_\gamma-\frac{1}{\gamma})$ for $\gamma<0$ and below by the same value for $\gamma>0$. Then, the second integral, say G, can be expressed as

$$G=\mathrm{sign}(\gamma)\int_{\bar{x}_{\gamma+\delta\gamma}-\frac{1}{\gamma+\delta\gamma}}^{\bar{x}_\gamma-\frac{1}{\gamma}}(1+(\gamma+\delta\gamma)(x-\bar{x}_{\gamma+\delta\gamma}))^v Q(x)\,\mathrm{d}x, \tag{B.3}$$

$$=\frac{\mathrm{sign}(\gamma)}{\gamma+\delta\gamma}\int_0^a y^v Q\left(\frac{y-1}{\gamma+\delta\gamma}+\bar{x}_{\gamma+\delta\gamma}\right)\mathrm{d}y \tag{B.4}$$

with $a=(\gamma+\delta\gamma)(\bar{x}_{\gamma+\delta\gamma}-\bar{x}_\gamma)-\frac{\delta\gamma}{\gamma}$, that tends to zero with $\delta\gamma$ if $\bar{x}_\gamma$ is continuous. At first order, we then obtain

$$G = \operatorname{sign}(\gamma)\frac{Q\left(\bar{x}_{\gamma+\delta\gamma} - \frac{1}{\gamma+\delta\gamma}\right)}{\gamma+\delta\gamma}\int_0^a y^v \mathrm{d}y \propto \frac{a^{1+v}}{1+v}$$

for $v > -1$. Then, it is readily checked that $\lim_{\delta\gamma\to 0}\frac{1}{\delta\gamma}G = 0$ for $v > 0$, so that (B.2) is always zero for $v > 0$ and (14) is true.

In the discrete case, the partition function is

$$Z_v(\gamma, \bar{x}_\gamma) = \sum_{x\in\mathscr{D}}(1+\gamma(x-\bar{x}_\gamma))^v Q(x).$$

There exists singular isolated values of γ such that $1+\gamma(x-\bar{x}_\gamma) = 0$, for x integer. For such values, the corresponding term in the partition function diverges for $v < 0$. Contrary to the continuous case where the domain of γ is contiguous, the domain of values of γ ensuring that the partition function is finite will be interrupted by isolated values of γ: the domain of possible γ will be constituted of segments. As in the continous case, the derivative of the partition function writes as the sum of two terms, the second one involving a domain increment

$$\frac{\mathrm{d}Z_v(\gamma,\bar{x}_\gamma)}{\mathrm{d}\gamma} = \sum_{\mathscr{D}}\frac{\mathrm{d}}{\mathrm{d}\gamma}(1+\gamma(x-\bar{x}_\gamma))^v Q(x) \tag{B.5}$$

$$+\lim_{\delta\gamma\to 0}\frac{1}{\delta\gamma}\sum_{\delta\mathscr{D}}(1+(\gamma+\delta\gamma)(x-\bar{x}_{\gamma+\delta\gamma}))^v Q(x). \tag{B.6}$$

If $\mathscr{D}$ does not depend on γ, there is no domain increment and the derivative is given by (B.5). When the bounds of $\mathscr{D}$ depend of γ, the domain increment is given by the integers in the interval $(\lceil \bar{x}_{\gamma+\delta\gamma} - \frac{1}{\gamma+\delta\gamma}\rceil, \lceil \bar{x}_\gamma - \frac{1}{\gamma}\rceil)$ $(\gamma > 0)$ or $(\lfloor \bar{x}_\gamma - \frac{1}{\gamma}\rfloor, \lfloor \bar{x}_{\gamma+\delta\gamma} - \frac{1}{\gamma+\delta\gamma}\rfloor)$ $(\gamma < 0)$; where $\lfloor x\rfloor$ is the floor function that returns the largest integer less than or equal to x; and $\lceil x\rceil$ is the ceil function, the smallest integer not less than x. If γ belongs in some interval such that the domain increment remains empty, then the derivative is of course simply (B.5). An extension will occur for an infinitesimal variation $\delta\gamma$ if $\bar{x}_\gamma - \frac{1}{\gamma}$ is precisely an integer, say k,

Then, the second sum reduces to

$$G = (1+(\gamma+\delta\gamma)(k-\bar{x}_{\gamma+\delta\gamma}))^v Q(k) \tag{B.7}$$

$$= \left(-\frac{\delta\gamma}{\gamma} - (\gamma+\delta\gamma)(\bar{x}_{\gamma+\delta\gamma} - \bar{x}_\gamma)\right)^v Q(k), \tag{B.8}$$

and finally

$$\lim_{\delta\gamma\to 0}\frac{1}{\delta\gamma}G = \lim_{\delta\gamma\to 0}\delta\gamma^{v-1}\left((\gamma+\delta\gamma)\frac{(\bar{x}_{\gamma+\delta\gamma} - \bar{x}_\gamma)}{\delta\gamma} - \frac{1}{\gamma}\right)^{1+v} = 0 \quad \text{for } v > 1 \tag{B.9}$$

since all terms in the parenthesis remains finite when $\delta\gamma \to 0$. In such case the derivative reduces to (B.5) and (14) is true.

References

[1] M. Asadi, I. Bayramoglu, The mean residual life function of a k-out-of-n structure at the system level, IEEE Transactions on Reliability 55 (2006) 314–318.

[2] A. Banerjee, S. Merugu, I.S. Dhillon, J. Ghosh, Clustering with Bregman divergences, Journal of Machine Learning Research 6 (2005) 1705–1749.

[3] R. Baraniuk, P. Flandrin, A. Janssen, O. Michel, Measuring time-frequency information content using the Rényi entropies, IEEE Transactions on Information Theory 47 (2001) 1391–1409.

[4] A.G. Bashkirov, On maximum entropy principle, superstatistics, power-law distribution and Rényi parameter, Physica A 340 (2004) 153–162.

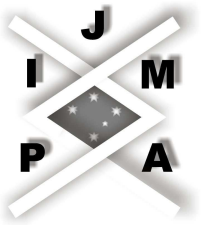
[5] M. Basseville, Distance measures for signal processing and pattern recognition, Signal Processing 18 (1989) 349–369.

[6] C. Beck, Generalized statistical mechanics of cosmic rays, Physica A 331 (2004) 173–181.

[7] D. Bhandari, N.R. Pal, Some new information measures for fuzzy sets, Information Sciences 67 (1993) 209–228.

[8] A.C. Cebrian, M. Denuit, P. Lambert, Generalized pareto fit to the society of actuaries' large claims database, North American Actuarial Journal 7 (2003) 18–36.
[9] I. Csiszár, Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems, Annals of Statistics 19 (1991) 2032–2066.
[10] I. Csiszár, Generalized cutoff rates and Rényi's information measures, IEEE Transactions on Information Theory 41 (1995) 26–34.
[11] R.S. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Grundlehren der mathematischen Wissenschaften, vol. 271, Springer-Verlag, 1985.
[12] M.D. Esteban, Divergence statistics based on entropy functions and stratified sampling, Information Sciences 87 (1995) 185–203.
[13] A. Golan, J.M. Perloff, Comparison of maximum entropy and higher-order entropy estimators, Journal of Econometrics 107 (2002) 195–211.
[14] M. Grendar, M. Grendar, Maximum entropy method with non-linear moment constraints: challenges, AIP Conference Proceedings (2004).
[15] Y. He, A. Hamza, H. Krim, A generalized divergence measure for robust image registration, IEEE Transactions on Signal Processing 51 (2003) 1211–1220, see also Acoustics, Speech, and Signal Processing.
[16] E.T. Jaynes, Information theory and statistical mechanics, Physical Review 108 (1957) 171.
[17] E.T. Jaynes, On the rationale of maximum entropy methods, Proceedings of the IEEE 70 (1982) 939–952.
[18] P. Jizba, T. Arimitsu, The world according to Rényi: thermodynamics of multifractal systems, Annals of Physics 312 (2004) 17–59.
[19] A. Krishnamachari, V. moy Mandal, Karmeshu, Study of dna binding sites using the Rényi parametric entropy measure, Journal of Theoretical Biology 227 (2004) 429–436.
[20] S. Kullback, Information Theory and Statistics, Wiley, New York, 1959.
[21] B. LaCour, Statistical characterization of active sonar reverberation using extreme value theory, IEEE Journal of Oceanic Engineering 29 (2004) 310–316.
[22] M.M. Mayoral, Rényi's entropy as an index of diversity in simple-stage cluster sampling, Information Sciences 105 (1998) 101–114.
[23] I. Molina, D. Morales, Rényi statistics for testing hypotheses in mixed linear regression models, Journal of Statistical Planning and Inference 137 (2007) 87–102.
[24] M.A.J.V. Montfort, J.V. Witter, Generalized Pareto distribution applied to rainfall depths, Hydrological Sciences Journal 31 (1986) 151–162.
[25] S. Nadarajah, K. Zografos, Formulas for Rényi information and related measures for univariate distributions, Information Sciences 155 (2003) 119–138.
[26] S. Nadarajah, K. Zografos, Expressions for Rényi and Shannon entropies for bivariate distributions, Information Sciences 170 (2005) 173–189.
[27] A.K. Nanda, S.S. Maiti, Rényi information measure for a used item, Information Sciences 177 (2007) 4161–4175.
[28] J. Naudts, Dual description of nonextensive ensembles, Chaos, Solitons, & Fractals 13 (2002) 445–450.
[29] H. Neemuchwala, A. Hero, P. Carson, Image matching using alpha-entropy measures and entropic graphs, Signal Processing 85 (2005) 277–296.
[30] R. Nock, F. Nielsen, On weighting clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (2006) 1223–1235.
[31] G.A. Raggio, On equivalence of thermostatistical formalisms, 1999. <http://arxiv.org/abs/cond-mat/9909161>.
[32] A. Rényi, On Measures of Entropy and Information, University of California Press, Berkeley, CA, 1961.
[33] K.-S. Song, Rényi information, loglikelihood and an intrinsic distribution measure, Journal of Statistical Planning and Inference 93 (2001) 51–69.
[34] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, Journal of Statistical Physics 52 (1988) 479–487.
[35] C. Tsallis, Entropic nonextensivity: a possible measure of complexity, Chaos, Solitons, & Fractals 13 (2002) 371–391.
[36] C. Tsallis, R.S. Mendes, A.R. Plastino, The role of constraints within generalized nonextensive statistics, Physica A 261 (1998) 534–554.
[37] C. Vignat, A. Hero, J.A. Costa, About closedness by convolution of the Tsallis maximizers, Physica A 340 (2004) 147–152.
[38] S. Vinga, J.S. Almeida, Rényi continuous entropy of DNA sequences, Journal of Theoretical Biology 231 (2004) 377–388.

8.6 C. Vignat and J.-F. Bercher, "On Fisher information inequalities and score functions in non-invertible linear systems," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 4, p. Article 71, 2003.



Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 4, Issue 4, Article 71, 2003

ON FISHER INFORMATION INEQUALITIES AND SCORE FUNCTIONS IN NON-INVERTIBLE LINEAR SYSTEMS

C. VIGNAT AND J.-F. BERCHER
E.E.C.S. UNIVERSITY OF MICHIGAN
1301 N. BEAL AVENUE
ANN ARBOR MI 48109, USA.
vignat@univ-mlv.fr
URL: http://www-syscom.univ-mlv.fr/ vignat/

ÉQUIPE SIGNAL ET INFORMATION
ESIEE AND UMLV
93 162 NOISY-LE-GRAND, FRANCE.
jf.bercher@esiee.fr

Received 04 June, 2003; accepted 09 September, 2003
Communicated by F. Hansen

ABSTRACT. In this note, we review score functions properties and discuss inequalities on the Fisher Information Matrix of a random vector subjected to linear non-invertible transformations. We give alternate derivations of results previously published in [6] and provide new interpretations of the cases of equality.

Key words and phrases: Fisher information, Non-invertible linear systems.

2000 *Mathematics Subject Classification.* 62B10, 93C05, 94A17.

1. INTRODUCTION

The Fisher information matrix J_X of a random vector X appears as a useful theoretic tool to describe the propagation of information through systems. For instance, it is directly involved in the derivation of the Entropy Power Inequality (EPI), that describes the evolution of the entropy of random vectors submitted to linear transformations. The first results about information transformation were given in the 60's by Blachman [1] and Stam [5]. Later, Papathanasiou [4] derived an important series of Fisher Information Inequalities (FII) with applications to characterization of normality. In [6], Zamir extended the FII to the case of non-invertible linear systems. However, the proofs given in his paper, completed in the technical report [7], involve complicated derivations, especially for the characterization of the cases of equality.

The main contributions of this note are threefold. First, we review some properties of score functions and characterize the estimation of a score function under linear constraint. Second,

ISSN (electronic): 1443-5756

Part of this work was presented at International Symposium on Information Theory, Lausanne, Switzerland, 2002.
077-03

we give two alternate derivations of Zamir's FII inequalities and show how they can be related to Papathanasiou's results. Third, we examine the cases of equality and give an interpretation that highlights the concept of extractable component of the input vector of a linear system, and its relationship with the concepts of pseudoinverse and gaussianity.

2. NOTATIONS AND DEFINITIONS

In this note, we consider a linear system with a $(n \times 1)$ random vector input X and a $(m \times 1)$ random vector output Y, represented by a $m \times n$ matrix A, with $m \leq n$ as

$$Y = AX.$$

Matrix A is assumed to have full row rank (rank $A = m$).

Let f_X and f_Y denote the probability densities of X and Y. The probability density f_X is supposed to satisfy the three regularity conditions (cf. [4])

(1) $f_X(x)$ is continuous and has continuous first and second order partial derivatives,
(2) $f_X(x)$ is defined on $\mathbb{R}^n$ and $\lim_{||x||\to\infty} f_X(x) = 0$,
(3) the Fisher information matrix J_X (with respect to a translation parameter) is defined as

$$[J_X]_{i,j} = \int_{\mathbb{R}^n} \left[\frac{\partial \ln f_X(x)}{\partial x_i} \frac{\partial \ln f_X(x)}{\partial x_j} \right] f_X(x) \mathrm{d}x,$$

and is supposed nonsingular.

We also define the score functions $\phi_X(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ associated with f_X according to:

$$\phi_X(x) = \frac{\partial \ln f_X(x)}{\partial x}.$$

The statistical expectation operator E_X is

$$\mathrm{E}_X\left[h(X)\right] = \int_{\mathbb{R}^n} h(x) f_X(x) \mathrm{d}x.$$

$\mathrm{E}_{X,Y}$ and $\mathrm{E}_{X|Y}$ will denote the mutual and conditional expectations, computed with the mutual and conditional probability density functions $f_{X,Y}$ and $f_{X|Y}$ respectively.

The covariance matrix of a random vector $g(X)$ is defined by

$$\mathrm{cov}[g(X)] = \mathrm{E}_X\left[(g(X) - \mathrm{E}_X\left[g(X)\right])(g(X) - \mathrm{E}_X\left[g(X)\right])^T\right].$$

The gradient operator ∇_X is defined by

$$\nabla_X h(X) = \left[\frac{\partial h(X)}{\partial x_1}, \ \ldots, \ \frac{\partial h(X)}{\partial x_n} \right]^T.$$

Finally, in what follows, a matrix inequality such as $A \geq B$ means that matrix $(A - B)$ is nonnegative definite.

3. PRELIMINARY RESULTS

We derive here a first theorem that extends Lemma 1 of [7]. The problem addressed is to find an estimator $\widehat{\phi_X(X)}$ of the score function $\phi_X(X)$ in terms of the observations $Y = AX$. Obviously, this estimator depends of Y, but this dependence is omitted here for notational convenience.

Theorem 3.1. *Under the hypotheses expressed in Section 2, the solution to the minimum mean square estimation problem*

$$\widehat{\phi_X(X)} = \arg\min_w \mathrm{E}_{X,Y}\left[||\phi_X(X) - w(Y)||^2\right] \text{ subject to } Y = AX, \tag{3.1}$$

is

$$\widehat{\phi_X(X)} = A^T \phi_Y(Y). \tag{3.2}$$

The proof we propose here relies on elementary algebraic manipulations according to the rules expressed in the following lemma.

Lemma 3.2. *If X and Y are two random vectors such that $Y = AX$, where A is a full row-rank matrix then for any smooth functions $g_1 : \mathbb{R}^m \to \mathbb{R}$, $g_2 : \mathbb{R}^n \to \mathbb{R}$, $h_1 : \mathbb{R}^n \to \mathbb{R}^n$, $h_2 : \mathbb{R}^m \to \mathbb{R}^m$,*

Rule 0
$$\mathrm{E}_X[g_1(AX)] = \mathrm{E}_Y[g_1(Y)]$$

Rule 1
$$\mathrm{E}_X[\phi_X(X) g_2(X)] = -\mathrm{E}_X[\nabla_X g_2(X)]$$

Rule 2
$$\mathrm{E}_X\left[\phi_X(X) h_1^T(X)\right] = -\mathrm{E}_X\left[\nabla_X h_1^T(X)\right]$$

Rule 3
$$\nabla_X h_2^T(AX) = A^T \nabla_Y h_2^T(Y)$$

Rule 4
$$\mathrm{E}_X\left[\nabla_X \phi_Y^T(Y)\right] = -A^T J_Y.$$

Proof. Rule 0 is proved in [2, vol. 2, p.133]. Rule 1 and Rule 2 are easily proved using integration by parts. For Rule 3, denote by h_k the k^{th} component of vector $h = h_2$, and remark that

$$\begin{aligned}\frac{\partial}{\partial x_j} h_k(AX) &= \sum_{i=1}^m \frac{\partial h_k(AX)}{\partial y_i}\frac{\partial y_i}{\partial x_j} \\ &= \sum_{i=1}^m A_{ij}\left[\nabla_Y h_k(Y)\right]_i \\ &= \left[A^T \nabla_Y h_k(Y)\right]_j.\end{aligned}$$

Now $h^T(Y) = \left[h_1^T(Y), \ldots, h_n^T(Y)\right]$ so that $\nabla_X h^T(Y) = \left[\nabla_X h_1^T(AX), \ldots, \nabla_X h_n^T(AX)\right] = A^T \nabla_Y h^T(Y)$.

Rule 4 can be deduced as follows:

$$\begin{aligned}&\mathrm{E}_X\left[\nabla_X \phi_Y^T(Y)\right] \\ &\quad \overset{\text{Rule 3}}{=} A^T \mathrm{E}_X\left[\nabla_Y \phi_Y^T(Y)\right] \\ &\quad \overset{\text{Rule 0}}{=} A^T \mathrm{E}_Y\left[\nabla_Y \phi_Y^T(Y)\right] \\ &\quad \overset{\text{Rule 2}}{=} -A^T \mathrm{E}_Y\left[\phi_Y(Y)\phi_Y(Y)^T\right].\end{aligned}$$

□

For the proof of Theorem 3.1, we will also need the following orthogonality result.

Lemma 3.3. *For all multivariate functions $h : \mathbb{R}^m \to \mathbb{R}^n$, $\widehat{\phi_X(X)} = A^T \phi_Y(Y)$ satisfies*

$$\mathrm{E}_{X,Y}\left(\phi_X(X) - \widehat{\phi_X(X)}\right)^T h(Y) = 0. \tag{3.3}$$

Proof. Expand into two terms and compute first term using the trace operator $tr(\cdot)$

$$\begin{aligned}\mathrm{E}_{X,Y}\left[\phi_X(X)^T h(Y)\right] &= tr\,\mathrm{E}_{X,Y}\left[\phi_X(X)\, h^T(Y)\right] \\ &\overset{\text{Rule 2, Rule 0}}{=} -tr\,\mathrm{E}_Y\left[\nabla_X h^T(Y)\right] \\ &\overset{\text{Rule 3}}{=} -tr\ A^T \mathrm{E}_Y\left[\nabla_Y h^T(Y)\right].\end{aligned}$$

Second term writes

$$\begin{aligned}\mathrm{E}_{X,Y}\left[\widehat{\phi_X(X)}^T h(Y)\right] &= tr\,\mathrm{E}_{X,Y}\left[\widehat{\phi_X(X)} h^T(Y)\right] \\ &= tr\,\mathrm{E}_Y\left[A^T\phi_Y(Y)\, h^T(Y)\right] \\ &= tr\, A^T \mathrm{E}_Y\left[\phi_Y(Y)\, h^T(Y)\right] \\ &\overset{\text{Rule 2}}{=} -tr\ A^T \mathrm{E}_Y\left[\nabla_Y h^T(Y)\right]\end{aligned}$$

thus the terms are equal. □

Using Lemma 3.2 and Lemma 3.3 we are now in a position to prove Theorem 3.1.

Proof of Theorem 3.1. From Lemma 3.3, we have

$$\begin{aligned}\mathrm{E}_{X,Y}\left[\left(\phi_X(X) - \widehat{\phi_X(X)}\right) h(Y)\right] &= \mathrm{E}_{X,Y}\left[\left(\phi_X(X) - A^T\phi_Y(Y)\right) h(Y)\right] \\ &= \mathrm{E}_Y\left[\mathrm{E}_{X|Y}\left[\left(\phi_X(X) - A^T\phi_Y(Y)\right) h(Y)\right]\right] = 0.\end{aligned}$$

Since this is true for all h, it means the inner expectation is null, so that

$$\mathrm{E}_{X|Y}\left[\phi_X(X)\right] = A^T\phi_Y(Y).$$

Hence, we deduce that the estimator $\widehat{\phi_X(X)} = A^T\phi_Y(Y)$ is nothing else but the conditional expectation of $\phi_X(X)$ given Y. Since it is well known (see [8] for instance) that the conditional expectation is the solution of the Minimum Mean Square Error (MMSE) estimation problem addressed in Theorem 3.1, the result follows. □

Theorem 3.1 not only restates Zamir's result in terms of an estimation problem, but also extends its conditions of application since our proof does not require, as in [7], the independence of the components of X.

4. FISHER INFORMATION MATRIX INEQUALITIES

As was shown by Zamir [6], the result of Theorem 3.1 may be used to derive the pair of Fisher Information Inequalities stated in the following theorem:

Theorem 4.1. *Under the assumptions of Theorem 3.1,*

$$J_X \geq A^T J_Y A \tag{4.1}$$

and

$$J_Y \leq \left(A J_X^{-1} A^T\right)^{-1}. \tag{4.2}$$

We exhibit here an extension and two alternate proofs of these results, that do not even rely on Theorem 3.1. The first proof relies on a classical matrix inequality combined with the algebraic properties of score functions as expressed by Rule 1 to Rule 4. The second (partial) proof is deduced as a particular case of results expressed by Papathanasiou [4].

The first proof we propose is based on the well-known result expressed in the following lemma.

Lemma 4.2. *If* $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *is a block symmetric non-negative matrix such that* D^{-1} *exists, then*

$$A - BD^{-1}C \geq 0,$$

with equality if and only if $rank(U) = \dim(D)$.

Proof. Consider the block $L\Delta M$ factorization [3] of matrix U :

$$U = \underbrace{\begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}}_{\Delta} \underbrace{\begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}}_{M^T}.$$

We remark that the symmetry of U implies that $L = M$ and thus

$$\Delta = L^{-1}UL^{-T}$$

so that Δ is a symmetric nonnegative definite matrix. Hence, all its principal minors are non-negative, and

$$A - BD^{-1}C \geq 0.$$

□

Using this matrix inequality, we can complete the proof of Theorem 4.1 by considering the two following $(m+n) \times (m+n)$ matrices

$$U_1 = \mathrm{E}\begin{bmatrix} \phi_X(X) \\ \phi_Y(Y) \end{bmatrix} \begin{bmatrix} \phi_X^T(X) & \phi_Y^T(Y) \end{bmatrix},$$

$$U_2 = \mathrm{E}\begin{bmatrix} \phi_Y(Y) \\ \phi_X(X) \end{bmatrix} \begin{bmatrix} \phi_Y^T(Y) & \phi_X^T(X) \end{bmatrix}.$$

For matrix U_1, we have, from Lemma 4.2

$$\begin{aligned} &\mathrm{E}_X\left[\phi_X(X)\,\phi_X^T(X)\right] \\ &\geq \mathrm{E}_{X,Y}\left[\phi_X(X)\,\phi_Y^T(Y)\right]\left(\mathrm{E}_Y\left[\phi_Y(Y)\,\phi_Y^T(Y)\right]\right)^{-1}\mathrm{E}_{X,Y}\left[\phi_Y(Y)\,\phi_X^T(X)\right]. \end{aligned} \tag{4.3}$$

Then, using the rules of Lemma 3.2, we can recognize that

$$\begin{aligned} \mathrm{E}_X\left[\phi_X(X)\,\phi_X^T(X)\right] &= J_X, \\ \mathrm{E}_Y\left[\phi_Y(Y)\,\phi_Y^T(Y)\right] &= J_Y, \\ \mathrm{E}_{X,Y}\left[\phi_X(X)\,\phi_Y^T(Y)\right] &= -\mathrm{E}_Y\left[\nabla\phi_Y^T(Y)\right] = A^T J_Y, \\ \mathrm{E}_{X,Y}\left[\phi_Y(Y)\,\phi_X^T(X)\right] &= \left(A^T J_Y\right)^T = J_Y A. \end{aligned}$$

Replacing these expressions in inequality (4.3), we deduce the first inequality (4.1).

Applying the result of Lemma 4.2 to matrix U_2 yields similarly

$$J_Y \geq J_Y^T A J_X^{-1} A^T J_Y.$$

Multiplying both on left and right by $J_Y^{-1} = \left(J_Y^{-1}\right)^T$ yields inequality (4.2).

Another proof of inequality (4.2) is now exhibited, as a consequence of a general result derived by Papathanasiou [4]. This result states as follows.

Theorem 4.3. *(Papathanasiou* [4]*) If $g(X)$ is a function $\mathbb{R}^n \to \mathbb{R}^m$ such that, $\forall i \in [1,m]$, $g_i(x)$ is differentiable and $var[g_i(X)] \leq \infty$, the covariance matrix $cov[g(X)]$ of $g(X)$ verifies:*

$$\mathrm{cov}[g(X)] \geq \mathrm{E}_X\left[\nabla^T g(X)\right] J_X^{-1} \mathrm{E}_X\left[\nabla g(X)\right].$$

Now, inequality (4.2) simply results from the choice $g(X) = \phi_Y(AX)$, since in this case $\mathrm{cov}[g(X)] = J_Y$ and $\mathrm{E}_X\left[\nabla^T g(X)\right] = -J_Y A$. Note that Papathanasiou's theorem does not allow us to retrieve inequality (4.1).

5. CASE OF EQUALITY IN MATRIX FII

We now explicit the cases of equality in both inequalities (4.1) and (4.2). Case of equality in inequality (4.2) was already characterized in [7] and introduces the notion of 'extractable components' of vector X. Our alternate proof also makes use of this notion and establishes a link with the pseudoinverse of matrix A.

Case of equality in inequality (4.1). The case of equality in inequality (4.1) is characterized by the following theorem.

Theorem 5.1. *Suppose that components X_i of X are mutually independent. Then equality holds in (4.1) if and only if matrix A possesses $(n-m)$ null columns or, equivalently, if A writes, up to a permutation of its column vectors*

$$A = [A_0 \mid 0_{m\times(n-m)}],$$

where A_0 is a $m \times m$ non-singular matrix.

Proof. According to the first proof of Theorem 4.1 and the case of equality in Lemma 4.2, equality holds in (4.1) if there exists a non-random matrix B and a non-random vector c such that

$$\phi_X(X) = B\phi_Y(Y) + c.$$

However, as random variables $\phi_X(X)$ and $\phi_Y(Y)$ have zero-mean, $\mathrm{E}_X[\phi(X)] = 0$, $\mathrm{E}_Y[\phi(Y)] = 0$, then necessarily $c = 0$. Moreover, applying Rule 2 and Rule 4 yields

$$\mathrm{E}_{X,Y}\left[\phi_X(X)\,\phi_Y(Y)^T\right] = A^T J_Y$$

on one side, and

$$\mathrm{E}_{X,Y}\left[\phi_X(X)\,\phi_Y(Y)^T\right] = BJ_Y$$

on the other side, so that finally $B = A^T$ and

$$\phi_X(X) = A^T\phi_Y(Y).$$

Now, since A has rank m, it can be written, up to a permutation of its columns, under the form

$$A = [A_0 \,|\, A_0 M\,],$$

where A_0 is an invertible $m \times m$ matrix, and M is an $m \times (n-m)$ matrix. Suppose $M \neq 0$ and express equivalently X as

$$X = \begin{bmatrix} X_0 \\ X_1 \end{bmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix}$$

so that

$$\begin{aligned} Y &= AX \\ &= A_0X_0 + A_0MX_1 \\ &= A_0\tilde{X}, \end{aligned}$$

with $\tilde{X} = X_0 + MX_1$. Since A_0 is square and invertible, it follows that

$$\phi_Y(Y) = A_0^{-T}\phi_{\tilde{X}}\left(\tilde{X}\right)$$

so that

$$\begin{aligned}
\phi_X &= A^T\phi_Y(Y) \\
&= A^T A_0^{-T}\phi_{\tilde{X}}\left(\tilde{X}\right) \\
&= \begin{bmatrix} A_0^T \\ M^T A_0^T \end{bmatrix} A_0^{-T}\phi_{\tilde{X}}\left(\tilde{X}\right) \\
&= \begin{bmatrix} I \\ M^T \end{bmatrix} \phi_{\tilde{X}}\left(\tilde{X}\right) \\
&= \begin{bmatrix} \phi_{\tilde{X}}\left(\tilde{X}\right) \\ M^T\phi_{\tilde{X}}\left(\tilde{X}\right) \end{bmatrix}.
\end{aligned}$$

As X has independent components, ϕ_X can be decomposed as

$$\phi_X = \begin{bmatrix} \phi_{X_0}(X_0) \\ \phi_{X_1}(X_1) \end{bmatrix}$$

so that finally

$$\begin{bmatrix} \phi_{X_0}(X_0) \\ \phi_{X_1}(X_1) \end{bmatrix} = \begin{bmatrix} \phi_{\tilde{X}}\left(\tilde{X}\right) \\ M^T\phi_{\tilde{X}}\left(\tilde{X}\right) \end{bmatrix},$$

from which we deduce that

$$\phi_{X_1}(X_1) = M^T\phi_{X_0}(X_0).$$

As X_0 and X_1 are independent, this is not possible unless $M = 0$, which is the equality condition expressed in Theorem 5.1.

Reciprocally, if these conditions are met, then obviously, equality is reached in inequality (4.1). □

Case of equality in inequality (4.2). Assuming that components of X are mutually independent, the case of equality in inequality (4.2) is characterized as follows:

Theorem 5.2. *Equality holds in inequality (4.2) if and only if each component X_i of X verifies at least one of the following conditions*

a) *X_i is Gaussian,*
b) *X_i can be recovered from the observation of $Y = AX$, i.e. X_i is 'extractable',*
c) *X_i corresponds to a null column of A.*

Proof. According to the (first) proof of inequality (4.2), equality holds, as previously, if and only if there exists a matrix C such that

$$\phi_Y(Y) = C\phi_X(X), \tag{5.1}$$

which implies that $J_Y = CJ_XC^t$. Then, as by assumption $J_Y^{-1} = AJ_X^{-1}A^t$, $C = J_YAJ_X^{-1}$ is such a matrix. Denoting $\tilde{\phi}_X(X) = J_X^{-1}\phi_X(X)$ and $\tilde{\phi}_Y(Y) = J_Y^{-1}\phi_Y(Y)$, equality (5.1) writes

$$\tilde{\phi}_Y(Y) = A\tilde{\phi}_X(X). \tag{5.2}$$

The rest of the proof relies on the following two well-known results:

- if X is Gaussian then equality holds in inequality (4.2),

- if A is a non singular square matrix, equality holds in inequality (4.2) irrespectively of X.

We thus need to isolate the 'invertible part' of matrix A. In this aim, we consider the pseudoinverse $A^{\#}$ of A and form the product $A^{\#}A$. This matrix writes, up to a permutation of rows and columns

$$A^{\#}A = \begin{bmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix},$$

where I is the $n_i \times n_i$ identity, M is a $n_{ni} \times n_{ni}$ matrix and $\mathbf{0}$ is a $n_z \times n_z$ matrix with $n_z = n - n_i - n_{ni}$ (i stands for invertible, n_i for not invertible and z for zero). Remark that n_z is exactly the number of null columns of A. Following [6, 7], n_i is the number of 'extractable' components, that is the number of components of X that can be deduced from the observation $Y = AX$. We provide here an alternate characterization of n_i as follows: the set of solutions of $Y = AX$ is an affine set

$$X = A^{\#}Y + (I - A^{\#}A)Z = X_0 + (I - A^{\#}A)Z,$$

where X_0 is the minimum norm solution of the linear system $Y = AX$ and Z is any vector. Thus, n_i is exactly the number of components shared by X and X_0.

The expression of $A^{\#}A$ allows us to express $\mathbb{R}^n$ as the direct sum $\mathbb{R}^n = \mathbb{R}^i \oplus \mathbb{R}^{ni} \oplus \mathbb{R}^z$, and to express accordingly X as $X = \left[X_i^T, X_{ni}^T, X_z^T\right]^T$. Then equality in inequality (4.2) can be studied separately in the three subspaces as follows:

(1) restricted to subspace $\mathbb{R}^i$, A is an invertible operator, and thus equality holds without condition,
(2) restricted to subspace $\mathbb{R}^{ni}$, equality (5.2) writes $M\tilde{\phi}(X_{ni}) = \tilde{\phi}(MX_{ni})$ that means that necessarily all components of X_{ni} are gaussian,
(3) restricted to subspace $\mathbb{R}^z$, equality holds without condition.

□

As a final note, remark that, although A is supposed full rank, $n_i \leq \mathrm{rank}A$. For instance, consider matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

for which $n_i = 1$ and $n_{ni} = 2$. This example shows that the notion of 'extractability' should not be confused with the invertibility restricted to a subspace. A is clearly invertible in the subspace $x_3 = 0$. However, such a subspace is irrelevant here since, as we deal with continuous random input vectors, X has a null probability to belong to this subspace.

ACKNOWLEDGMENT

The authors wish to acknowledge Pr. Alfred O. Hero III at EECS for useful discussions and suggestions, particularly regarding Theorem 3.1. The authors also wish to thank the referee for his careful reading of the manuscript that helped to improve its quality.

REFERENCES

[1] N.M. BLACHMAN, The convolution inequality for entropy powers, *IEEE Trans. on Information Theory, IT* , **11** (1965), 267–271.

[2] W. FELLER, *Introduction to Probability Theory and its Applications*, New York, John Wiley & Sons, 1971.

[3] G. GOLUB, *Matrix Computations*, Johns Hopkins University Press, 1996.

[4] V. PAPATHANASIOU, Some characteristic properties of the Fisher information matrix via Cacoullos-type inequalities, *J. Multivariate Analysis*, **14** (1993), 256–265.

[5] A.J. STAM, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inform. Control*, **2** (1959), 101–112.

[6] R. ZAMIR, A proof of the Fisher information matrix inequality via a data processing argument, *IEEE Trans. on Information Theory, IT*, **44**(3) (1998), 1246–1250.

[7] R. ZAMIR, A necessary and sufficient condition for equality in the Matrix Fisher Information inequality, *Technical report, Tel Aviv University, Dept. Elec. Eng. Syst.*, 1997. Available online `http://www.eng.tau.ac.il/~zamir/techreport/crb.ps.gz`

[8] H.L. van TREES, *Detection, Estimation, and Modulation Theory*, Part I, New York London, John Wiley and Sons, 1968.

8.7 C. Vignat and J. F. Bercher, "Analysis of signals in the Fisher-Shannon information plane," *Physics Letters A*, vol. 312, no. 1-2, pp. 27–33, June 2003.

ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Physics Letters A 312 (2003) 27–33

PHYSICS LETTERS A

www.elsevier.com/locate/pla

Analysis of signals in the Fisher–Shannon information plane

C. Vignat [a,b,*], J.-F. Bercher [b]

[a] *EECS, University of Michigan, Ann Arbor, MI, USA*
[b] *Équipe Signal et Information, ESIEE and UMLV, France*

Received 21 February 2003; accepted 26 March 2003

Communicated by C.R. Doering

Abstract

We show that the analysis of complex, possibly non-stationary signals, can be carried out in an information plane, defined by both Shannon entropy and Fisher information. Our study is exemplified by two large families of distributions with physical relevance: the Student-t and the power exponentials.

PACS: 02.50.Cw; 02.50.Ey; 65.40.Gr

Keywords: Complexity; Fisher information; Shannon–Boltzmann entropy; Power exponential; Student-t distribution

1. Introduction

Fisher Information Measure (FIM) was introduced by Fisher in 1925 [9] in the context of statistical estimation. In the last ten years, a growing interest for this information measure has arisen in theoretical physics. In a seminal paper [1], Frieden has characterized FIM as a versatile tool to describe the evolution laws of physical systems; one of his major results is that the classical evolution equations (e.g., the Schrödinger wave equation, the Klein–Gordon equation, the Helmoltz wave equation, the diffusion equation, the Boltzmann and Maxwell–Boltzmann law) can be derived from the minimization of FIM under proper constraint. As another important result, a version of the H-theorem has been extended—under the name I-theorem—to the notion of Fisher information [3]. Until recently, Shannon (Boltzmann) entropy was considered as the major tool to describe the informational behavior and complexity of physical systems. The theoretical contributions cited above suggest that this judgment should be revised and that FIM appears as an appealing alternative to Shannon entropy.

Since FIM allows an accurate description of the behavior of dynamic systems, its application to the characterization of complex signals issued from these systems appears quite natural. This approach was adopted for example by Martin et al. [2] for the characterization of EEG signals. One of the interesting results of their study was that FIM allowed the detection of some non-stationary behavior in situations where the Shannon entropy shows limited dynamics.

* Corresponding author. Université de Marne-la-Vallée, Cité Descartes, 77454 Champs-sur-Marne cedex 2, France.
E-mail addresses: vignat@univ-mlv.fr (C. Vignat), bercherj@esiee.fr (J.-F. Bercher).

0375-9601/03/$ – see front matter
doi:10.1016/S0375-9601(03)00570-X

Motivated by this work, we define a Fisher–Shannon information plane and show that the simultaneous examination of both Shannon entropy and FIM may be required to characterize the non-stationary behavior of a complex signal. More precisely, we exhibit two families of physically relevant signals, the Tsallis signals and the power exponential signals which have the unexpected property that their temporal trajectory in the Fisher–Shannon (FS) information plane can be arbitrarily designed. As a consequence, any non-stationarity measurement device based on only one of these two measures would yield a suboptimal result.

This Letter is organized as follows: in Section 2, we review some basic properties of entropy and Fisher information measures and introduce the notion of the Fisher–Shannon plane. In Section 3, we study two families of parameterized random variables and their locations in the Fisher–Shannon plane. These results are used in Section 4 to build explicitly two families of signals whose trajectories in the FS plane can be designed arbitrarily.

2. Fisher's information measure and Shannon entropy power

In the following, we consider a random variable X whose probability density function is denoted as $f_X(x)$. Its Shannon entropy writes

$$H_X = -\int f_X(x)\log f_X(x)\,dx \tag{1}$$

and its Fisher information measure writes

$$I_X = \int \left(\frac{\partial}{\partial x} f_X(x)\right)^2 \frac{dx}{f_X(x)}. \tag{2}$$

For convenience, we will use, rather than entropy, the alternative notion of entropy power (see [4]), defined by

$$N_X = \frac{1}{2\pi e} e^{2H_X}. \tag{3}$$

Among other properties, both measures N_X and I_X verify a set of resembling inequalities.

- *Evolution law*: if X represents the state of a system, the second law of the thermodynamics writes

  $$\delta N_X \leqslant 0 \tag{4}$$

  whereas the I-theorem writes accordingly

  $$\delta I_X \leqslant 0. \tag{5}$$

- *Superadditivity property*: the Shannon entropy of the sum of two independent random variables verifies the entropy power inequality [4]

  $$N_{X+Y} \geqslant N_X + N_Y \tag{6}$$

  and the corresponding Fisher information inequality [4] writes

  $$I_{X+Y}^{-1} \geqslant I_X^{-1} + I_Y^{-1}. \tag{7}$$

Two additional properties will help us to track the information trajectory of a random signal.

- *The scaling property*: when scaling a random variable X by a scalar factor $a \in \mathbb{C}^*$, the entropy power and FIM transform as follows

  $$N_{aX} = |a|^2 N_X,$$
  $$I_{aX}^{-1} = |a|^2 I_X^{-1}.$$

- *The uncertainty property*:

  $$N_X I_X \geqslant 1 \tag{8}$$

  with equality if and only if X is a Gaussian random variable. A proof of this property can be found, for example, in [4].

Both scaling and uncertainty properties enlight the fact that FIM and Shannon entropy are intrinsically linked, so that the characterization of signals should be improved when considering their location in the Fisher–Shannon (FS) plane.

Definition 1. The FS area, denoted as $\mathcal{D}$, is the set of all reachable values of FIM and Shannon entropies, namely

$$\mathcal{D} = \big\{(N, I) \mid N \geqslant 0,\ I \geqslant 0 \text{ and } NI \geqslant 1\big\}. \tag{9}$$

Next, as a consequence of the scaling property, we note that a scaled version aX of a random variable X

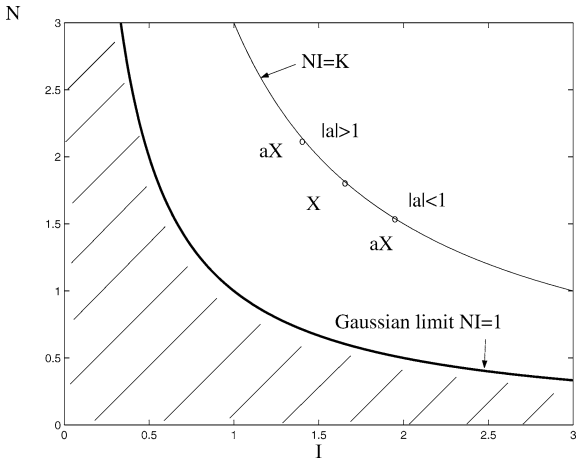


Fig. 1. The FS area.

belongs to the same $NI = K$ curve as X, as illustrated on Fig. 1.

3. Two families of random variables

In this part, we study two families of probability densities, namely the Student-t and the power exponential distributions. Both families include heavy tailed densities and the Student-t family is a member of the larger class of power laws. Furthermore, the Gaussian distribution is a particular case of both families and the uniform and exponential distributions are special cases of the power exponential family.

For both types of distributions, we compute the Shannon entropy power and FIM and the area spanned by these distributions in the FS plane.

3.1. Student-t random variable

The importance of Student-t distributions in statistical physics has been highlighted in the seminal work of Tsallis (see [7] and references therein). A part of the versatility of these distributions is due to the fact that they maximize under variance constraint a relevant statistical quantity, namely the Tsallis entropy

$$H_X^{(q)} = \frac{1}{q-1}\left(1 - \int f_X^q(x)\,dx\right), \tag{10}$$

where $q > 0$ is the extensivity parameter. Their exact expression is the following:

$$f_S(m,\sigma,x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sigma\sqrt{m-2}\,\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{1}{2}\right)} \times \left(1 + \frac{x^2}{(m-2)\sigma^2}\right)^{-\frac{m+1}{2}}$$
$$\text{for } (m > 2, 1 > q > 0, x \in \mathbb{R}) \tag{11}$$

where $m = \frac{1+q}{1-q}$ and σ is a scale parameter. These distributions and their properties have been extensively characterized, including in the multivariate case, in [6]. Note that the Gaussian law can be recovered from the Student-t family by letting $m \to +\infty$ or $q \to 1^-$. For convenience, we cite here two properties of these distributions that are helpful for the numerical simulations considered in Section 4.

- If X is distributed according to a Student-t law $f_S(m,\sigma,x)$, then a stochastic representation of X writes
$$X = \frac{\mathcal{N}}{\chi_m}, \tag{12}$$
where χ_m is a gamma distributed random variable with parameter m, independent of $\mathcal{N}$ which is Gaussian with zero mean and variance $\sigma^2(m-2)$.
- Moreover, if m is integer, χ_m is distributed as a chi variable with m degrees of freedom and a stochastic representation of X is thus
$$X = \frac{\mathcal{N}}{\sqrt{\sum_{k=1}^m \mathcal{N}_k^2}}, \tag{13}$$
where the $\{\mathcal{N}_k\}_{1\leqslant k\leqslant m}$ are independent Gaussian random variables with unit variance.

The information measures of a Student-t law are characterized in the following theorem.

Theorem 1. *The entropy power and the Fisher information measure of the Student-t distribution with parameters m and σ are*

$$N_S(m,\sigma) = \frac{1}{2\pi e}\left(\frac{\sigma\sqrt{m-2}\,\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)}\right)^2 \times e^{(m+1)\left(\psi\left(\frac{m+1}{2}\right)-\psi\left(\frac{m}{2}\right)\right)}, \tag{14}$$

$$I_S(m,\sigma) = \frac{1}{\sigma^2}\frac{m(m+1)}{(m-2)(m+3)}, \tag{15}$$

where ψ denotes the digamma function.

Proof. By direct computation. Result (14) has been already obtained, for example, in [8]. □

We are now in position to determine the area spanned by all Student-t laws in the Fisher–Shannon plane.

Theorem 2. *Denote $\mathcal{D}_S$ the area of the FS plane defined by*

$$\begin{cases} 1 \leqslant I_S N_S \leqslant \frac{3e^5}{80\pi}, \\ I_S \geqslant 0, \quad N_S \geqslant 0. \end{cases} \tag{16}$$

Then the application

$$]2,+\infty[\times[0,+\infty[\to \mathcal{D}_S, \quad (m,\sigma) \mapsto (I_S, N_S) \tag{17}$$

is a bijection.

Proof. The product $I_S(m,\sigma)N_S(m,\sigma)$ depends only on m since σ is scale parameter. Denote $h_S(m)$ this function so that

$$h_S(m) = \frac{m(m+1)}{2e(m+3)}\frac{\Gamma^2\left(\frac{m}{2}\right)}{\Gamma^2\left(\frac{m+1}{2}\right)} \times e^{(m+1)(\psi(\frac{m+1}{2})-\psi(\frac{m}{2}))} \quad (m>2). \tag{18}$$

It is easy to check that $h_S(m)$ decreases from $+\infty$ to 1 for $2 < m < +\infty$ and that $\lim_{m\to+\infty} h_S(m) = 1$ and $h_S(2) = \frac{3e^5}{80\pi}$. Thus the equation

$$I_S N_S = h_S(m) \tag{19}$$

determines uniquely $2 \leqslant m \leqslant +\infty$. Then, solving (15) in variable σ^2 yields a unique positive value of σ^2. □

3.2. Power exponential random variables

Let us now consider a second kind of random variable denoted as $X_{\mathrm{PE}}(\lambda,\gamma)$ whose probability density is of the power exponential type:

$$f_{\mathrm{PE}}(\lambda,\gamma,x) = \frac{\gamma\lambda^{1/\gamma}}{2\Gamma\left(\frac{1}{\gamma}\right)}\exp\left(-\lambda|x|^\gamma\right) \quad x\in\mathbb{R}, \tag{20}$$

where $\lambda > 0$ is a scale parameter and $\gamma > 1$ is a shape parameter. Note that $\gamma = 2$ gives the Gaussian case while $\gamma \to +\infty$ leads to a uniform distribution. This versatility explains that these distributions are used to model accurately some physically realistic quantities, for instance the amplitudes of wavelets coefficients. These distributions can also be shown to maximize the Shannon entropy under moment constraint [10].

The information measures associated with this probability density function are characterized in the following theorem.

Theorem 3. *The entropy power and the FIM of the power exponential law with parameters λ and γ are*

$$N_{\mathrm{PE}}(\lambda,\gamma) = \frac{2}{2\pi e}\left(\frac{1}{\gamma}\Gamma\left(\frac{1}{\gamma}\right)\left(\frac{e}{\lambda}\right)^{1/\gamma}\right)^2, \tag{21}$$

$$I_{\mathrm{PE}}(\lambda,\gamma) = \frac{\Gamma\left(1-\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)}\gamma(\gamma-1)\lambda^{2/\gamma}. \tag{22}$$

Proof. By direct computation. Result (21) has already been derived in [5]. □

Now given any point (I,N) in the Fisher–Shannon area, the following theorem proves that an exponential power random variable $X_{\mathrm{PE}}(\lambda,\gamma)$ having I and N as respective FIM and entropy power can be found.

Theorem 4. *Both applications*

$$]0,+\infty[\times]2,+\infty[\to \mathcal{D}, \quad (\lambda,\gamma)\mapsto(I,N) \tag{23}$$

and

$$]0,+\infty[\times]1,2[\to \mathcal{D}, \quad (\lambda,\gamma)\mapsto(I,N) \tag{24}$$

are bijections.

Proof. Suppose that we are given a couple $(I,N)\in\mathcal{D}$ and we want to determine (λ,γ). As λ is a scale parameter, the product $N_{\mathrm{PE}}(\lambda,\gamma)I_{\mathrm{PE}}(\lambda,\gamma) \geqslant 1$ is a function of the shape parameter γ only: denote $h_{\mathrm{PE}}(\gamma)$ this function so that

$$h_{\mathrm{PE}}(\gamma) = \frac{2e^{\frac{2}{\gamma}-1}}{\pi}\Gamma\left(\frac{1}{\gamma}\right)\Gamma\left(2-\frac{1}{\gamma}\right). \tag{25}$$

A straightforward computation shows that function h_{PE} decreases from $+\infty$ to 1 for $1 < \gamma \leqslant 2$ and increases from 1 to $+\infty$ for $2 \leqslant \gamma \leqslant +\infty$. Thus equation $h_{\mathrm{PE}}(\gamma) = NI$ has two solutions γ_1 and γ_2 such that $1 < \gamma_1 \leqslant 2$ and $2 \leqslant \gamma_2$. Given one of these solutions, λ can be determined uniquely by solving equation (21) or (22). □

4. Application to signal analysis

We are now in position to build non-stationary signals having arbitrary time trajectories in the FS plane. An an example, we have designed explicitly two signals:

- A first signal, based on independent Student-t samples, whose entropy power and FIM describe through time a step-like trajectory in the domain $\mathcal{D}_S$ of the FS plane.
- A second signal, based on independent power exponential samples, whose entropy and FIM evolve through time along a circle in the domain $\mathcal{D}$ of the FS plane.

4.1. The steps trajectory

The signal $s_{\mathrm{steps}}(n)$ is built from 112 000 independent samples of a Student-t with non-stationary parameters $m(n)$ and $\sigma(n)$ chosen as follows:

- Over the first 28 000 samples (A to B), parameter $m(n)$ increases linearly from $m = 3$ to $m = 30$, of one unit each 1000 samples; parameter $\sigma(n)$ evolves according to the following rule

$$\sigma^2(n) = I_S\big(m(n), 1\big) \tag{26}$$

  so that the FIM of $s_{\mathrm{steps}}(n)$ remains equal to 1 over all samples.
- Over the next 28 000 samples (B to C), parameter $m(n)$ decreases linearly from $m = 30$ to $m = 3$, of one unit each 1000 samples; parameter $\sigma(n)$ evolves according to the following rule

$$\sigma(n) = \frac{1}{\sqrt{N_S(m(n), 1)}} \tag{27}$$

  so that the entropy power remains equal to 1 over all samples.
- These two first steps are repeated accordingly over the 56 000 remaining samples so that the final trajectory in the FS plane looks as shown on Fig. 2: the first 28 000 of them (C to D) have a linearly decreasing entropy power with FIM fixed to 1.3 and conversely for the last part of the signal (D to E) where the entropy power is fixed to $N = 0.75$. Fig. 3 shows the corresponding signal $s_{\mathrm{steps}}(n)$.

This example shows that any measurement based on one only of the two information measures would yield a suboptimal detection.

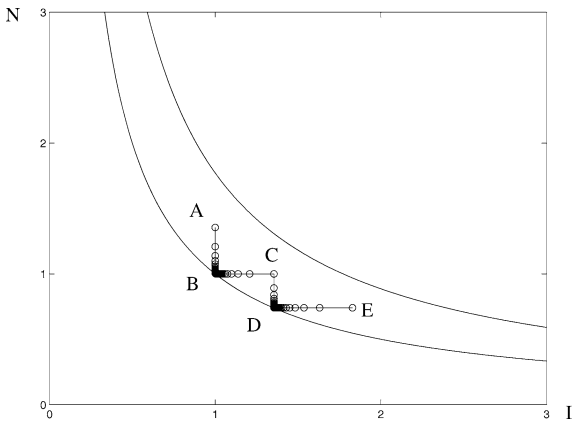


Fig. 2. The step-like trajectory of $s_{\mathrm{steps}}(n)$.

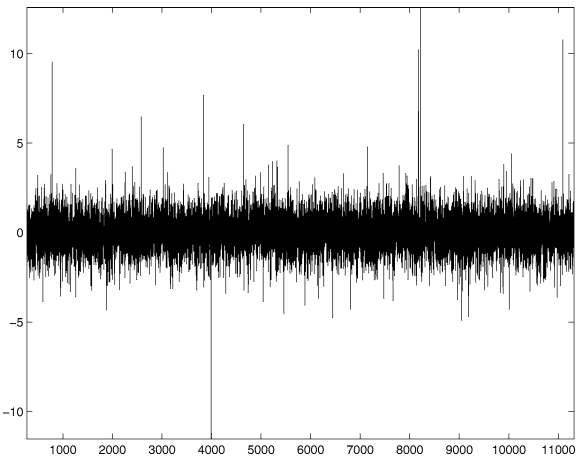


Fig. 3. Signal $s_{\mathrm{steps}}(n)$.

4.2. The circle trajectory

A signal was generated by drawing independent samples of power exponential distribution such as their entropy powers and FIM describe the circle trajectory in the FS plane as depicted on Fig. 4: to each point of the circle correspond 1000 samples, the circle being described clock-wise starting from angle $\theta = 0$, and step $\Delta\theta = \pi/10$. Details of the whole signal, corresponding to angles $\theta = \pi/5$ and $\theta = \pi + \pi/5$, are shown on Fig. 5.

This example shows once again that any trajectory can be designed in the FS plane and that a non-stationarity tracking device should inspect both FIM and entropy power to provide a good characterization of the signal.

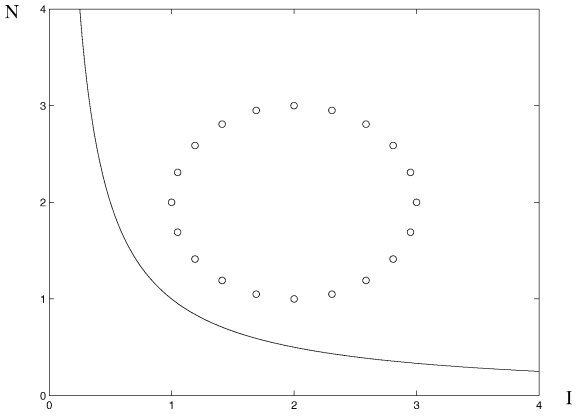


Fig. 4. Trajectory of $s_{\text{circle}}(n)$.

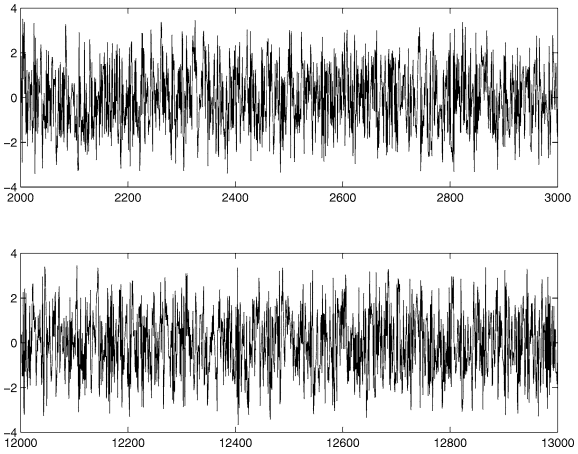


Fig. 5. Signal $s_{\text{circle}}(n)$.

4.3. Variance analysis

Given any point (I, N) in the area $\mathcal{D}_S$ of the FS plane, it follows from Theorems 2 and 4 that there exists a unique Student-t and a unique power exponential distribution having (I, N) as coordinates. But the variances of these distributions may also differ or not.

Fig. 6 shows constant variance trajectories of Student-t and power exponentials distributions: curve (1) (respectively (2)) corresponds to the family of Student-t (respectively power exponentials) with variance equal to 0.4, while curve (3) (respectively (4)) corresponds to a variance of 0.3. We deduce from this graph that there exist Student-t and power exponential laws that may be distinguished (separated) in the FS plane, although their variances coincide. Hence, in a non-stationary context, analysis in the FS plane may be useful.

On Fig. 7, a signal with a constant variance $\sigma^2 = 2$ is presented: first and last 2000 samples were generated according to a power exponential with parameters $(\lambda, \gamma) = (0.25, 2.05)$ and coordinates $(0.975, 1.74)$ in the FS plane; the 2000 center samples were generated according to a Student-t with parameters $(m, \sigma^2) = (3.5, 2)$ and coordinates $(0.807, 1.567)$.

Conversely, curves (5) and (6) in Fig. 6 correspond respectively to the trajectories of all Student-t distributions with variance equal to 36 and to all power expo-

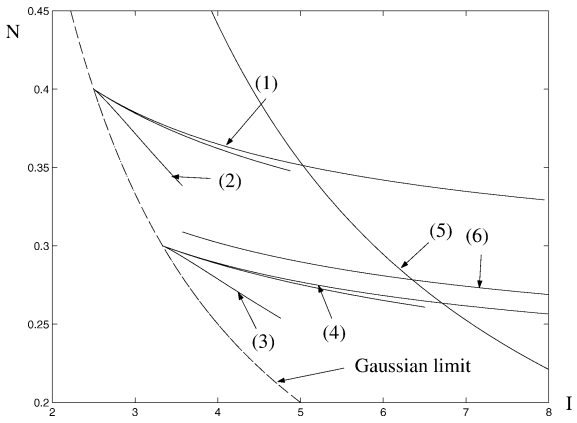


Fig. 6. Variance study in the FS plane.

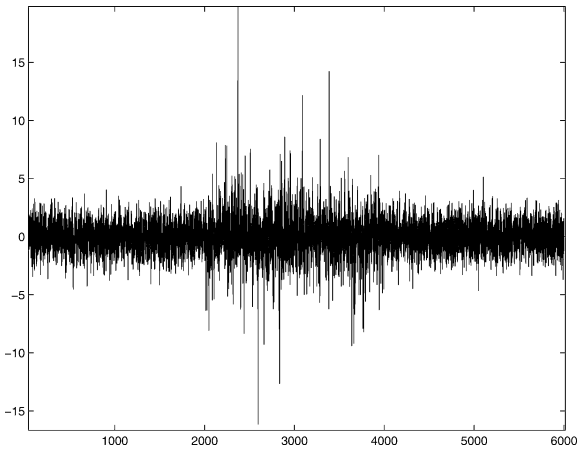


Fig. 7. Signal $s(n)$ with constant variance but different FS locations.

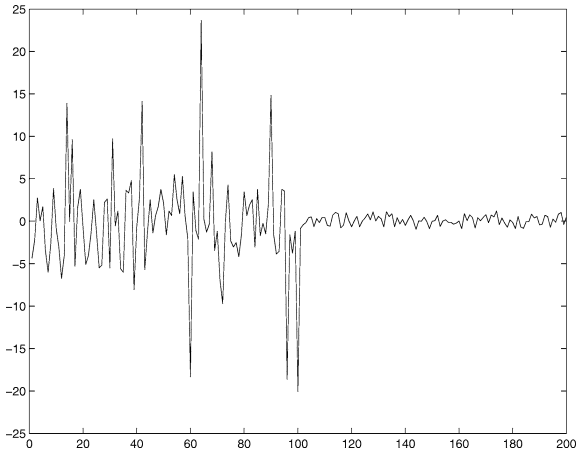


Fig. 8. Signal $s(n)$ generated according to distributions with same FS location but different variances.

nentials with variance 0.317. The intersection of these two curves correspond to a Student with parameters $m = 2.005$ and $\sigma = 6$ and to a power exponential with parameters $\lambda = 1$ and $\gamma = 6.25$. These two distributions are thus indistinguishable in the FS plane, while their variances are very different. A signal $s(n)$ generated from 1000 independent samples of the first distribution followed by 1000 samples of the second distribution, is shown on Fig. 8.

From this study, we conclude that in order to capture the dynamics of the complex signal, the additional information provided by variance may be useful, resulting in the analysis in the extended space (σ^2, I, N).

5. Conclusion

In this Letter, we have shown that two information measures, the Fisher information measure and the Shannon entropy power, can be used jointly in a context of non-stationarity detection. We have exhibited families of relevant signals that can behave arbitrarily in the Fisher–Shannon plane, proving that a device based on only one of these measures may fail to track the non-stationarity of the signal. We insist on the fact that the Student-t and power exponential distributions used here are not trivial examples, since both families span a large set of classical distributions and as such are often used to model complex data.

We can conclude that the EEG signals exhibited in [2] have a nearly horizontal trajectory in the FS plane and thus are better described by the inspection of their Fisher information rather by their entropy (power). We are currently working on the interpretation of such a property (and of its converse) for real data and trying to check its versatility for a larger class of physically relevant phenomena. We are also studying the geometrical characterization of signals in the FS plane, in terms of information distances and projections.

Acknowledgement

The authors wish to thank Dr. Bryan Griffiths who helped correcting the manuscript.

References

[1] B.R. Frieden, Phys. Rev. A 41 (1990) 4265.
[2] M.T. Martin, F. Pennini, A. Plastino, Phys. Lett. A 256 (1999) 173.
[3] A.R. Plastino, M. Casas, A. Plastino, Phys. Lett. A 246 (1998) 498.
[4] A. Dembo, T.M. Cover, J.A. Thomas, IEEE Trans. Inform. Theory 37 (1991) 1501.
[5] A. Kagan, Z. Landsman, Statist. Probab. Lett. 42 (1999) 7.
[6] J.A. Costa, A.O. Hero, C. Vignat, in: IEEE International Symposium on Information Theory, ISIT, Lausanne, 2002, p. 263.
[7] A.M.C. de Souza, C. Tsallis, Physica A 236 (1997) 52.
[8] K. Zografos, J. Multivariate Anal. 71 (1999) 67.
[9] R.A. Fisher, Proc. Cambridge Philos. Soc. 22 (1925) 700.
[10] C. Vignat, J.-F. Bercher, ESIEE Internal report number SIGTEL-SI 2003-2, 2003.

8.8 J.-F. Bercher and C. Vignat, "On minimum Fisher information distributions with restricted support and fixed variance," *Information Sciences*, vol. 179, pp. 3832-3842, Nov. 2009.

Information Sciences 179 (2009) 3832–3842

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On minimum Fisher information distributions with restricted support and fixed variance

J.-F. Bercher [a,*], C. Vignat [b]

[a] Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France
[b] Université de Marne-la-Vallée, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France

ARTICLE INFO

Article history:
Received 2 March 2009
Received in revised form 23 June 2009
Accepted 27 July 2009

Keywords:
Fisher information
Minimization problems
Whittaker functions
Differential equations

ABSTRACT

Fisher information is of key importance in estimation theory. It also serves in inference problems as well as in the interpretation of many physical processes. The mean-squared estimation error for the location parameter of a distribution is bounded by the inverse of the Fisher information associated with this distribution. In this paper we look for minimum Fisher information distributions with a restricted support. More precisely, we study the problem of minimizing the Fisher information in the set of distributions with fixed variance defined on a bounded subset $\mathscr{S}$ of $\mathbb{R}$ or on the positive real line. We show that the solutions of the underlying differential equation can be expressed in terms of Whittaker functions. Then, in the two considered cases, we derive the explicit expressions of the solutions and investigate their behavior. We also characterize the behavior of the minimum Fisher information as a function of the imposed variance.

1. Introduction

Importance of Fisher information as a measure of the information in a distribution is well known. It has many implications in estimation theory, as exemplified by the Cramér-Rao bound which is a fundamental limit on the variance of an estimator. Recent applications for computing performance bounds can be found in [2,35]. It is used as a method of inference and understanding in statistical physics and biology, as promoted by Frieden [10–12]. It is also used as a tool for characterizing complex signals or systems, [17,20,30] with applications, e.g. in geophysics [25,23,3,16], in biology [9], in reconstruction [5,6,21] or in signal processing [38,37,29]. Other applications are in random censoring [26], hypothesis testing [19], classification [8]. In robust estimation, minimization of Fisher information has been originally considered by Huber [15], and in the case of scale and location parameters in a Kolmogorov neighborhood of a parent distribution in [33,34]. Fisher information for orthogonal polynomials and special functions have been studied in [24,36]. Connections with the differential equations of Physics have been explored in [11]. It is also interesting to mention that Fisher information associated to a distribution appears in quantum physics under the name of Weiszäcker energy; in this setting, several inequalities for the Fisher information are derived in [22].

It is well known that the distribution with a fixed variance that minimizes the Fisher information on $\mathbb{R}$ is the standard gaussian distribution. However, there are many situations where the variables at hand are known to belong to some subset of $\mathbb{R}$. For instance, the random variable may be known, on physical grounds, to have only non negative outcomes, e.g. the variable represents an energy. Variables may also be known to have a distribution with a support restricted to a given interval: this is the case of normalized variables or of the measurements obtained from a physical device with a (necessarily)

* Corresponding author.
E-mail addresses: jf.bercher@esiee.fr (J.-F. Bercher), vignat@univ-mlv.fr (C. Vignat).

0020-0255/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2009.07.013

finite output range. In such cases, a model of the data probability distribution can be defined by the distribution compatible with the constraints extracted from the data with minimum (Fisher) information. This idea, advocated by Frieden and others, is reminiscent of the idea of maximum entropy distributions, when the measure of information considered is the Fisher information. By the Cramér-Rao bound, the Fisher information serves as a benchmark for estimators.

In the case of the estimation of a location parameter, the distributions with minimum Fisher information correspond to the most difficult estimation cases, and thus it is interesting to look for these distributions. This problem has been considered in the important paper [28], which reading has motivated the present work. This paper presents general results and characterizations of the solution to the minimization of Fisher information on a compact support, subject to a variance constraint. Here, we give the explicit closed-form expressions of the solutions, characterize their behavior and propose alternate simpler proofs. We also extend the results to the case of distributions with support confined to the positive half line (and of course the case of any other semi-infinite support as a straightforward consequence).

Let f denote the probability density of a random variable X. The Fisher information (with respect to a translation parameter) is defined as

$$I[f] = \int_{\mathscr{S}} \left(\frac{d\ln f(x)}{dx}\right)^2 f(x)\mathrm{d}x = \int_{\mathscr{S}} \left(\frac{df(x)}{dx}\right)^2 \frac{1}{f(x)}\mathrm{d}x, \tag{1}$$

where $\mathscr{S} = \mathrm{Supp}[f]$ denotes the support of the density, $f(x)$ is supposed differentiable and both $f(x)$ and its derivative $f'(x)$ are square integrable on $\mathscr{S}$. We note $\mathscr{D}$ the set of functions that verify these hypotheses. It is known [7] that the Fisher information is a strictly convex function of the distribution, that is: for $f(x), g(x)$, and with $\lambda \in (0,1)$, then

$$I[\lambda f + (1-\lambda)g] < \lambda I[f] + (1-\lambda)I[g]. \tag{2}$$

We are here interested in classifying distributions with a given variance, and consider the variational problem

$$I(\sigma^2) = \inf_f \{I[f] : \mathrm{Supp}[f] = \mathscr{S}, \quad f \in \mathscr{D} \text{ and } \mathrm{Var}[f] = \sigma^2\} \tag{3}$$

which consists in finding a distribution with minimum Fisher information on the set of all distributions with support $\mathscr{S}$ and a fixed variance σ^2. The value of the minimum Fisher information obtained for a given variance σ^2 is denoted $I(\sigma^2)$ – the use of the square brackets and parenthesis distinguishes between the functionals of the probability distributions and the functionals of the variance. Although $I[f]$ is a convex functional, the set defined by the constraint $\mathrm{Var}[f] = \sigma^2$ is not convex, so that uniqueness of the solution is not guaranteed. For instance, on $\mathbb{R}$, the normal distribution minimizes Fisher information in the set of distributions on $\mathbb{R}$ with a given variance, but in fact irrespectively of the value of the mean: all normal distributions, whatever their mean, are equivalent solutions. We will obtain in the following that solutions on $\mathbb{R}^+$ or on an interval are in fact unique.

In Section 2 we give the differential equation that is associated with the problem of minimization of Fisher information, and then we underline its relationship with known differential equations. So doing, we exhibit some explicit expressions of the solutions, in terms of Whittaker and parabolic cylinder functions. Then, we examine two particular cases. First, in Section 4 we characterize the solutions with positive support, the behavior of the minimum Fisher information with respect to the variance, and show that the solution is unique and turns out to be a chi distribution. Second, in Section 5 where the support is restricted to an interval, we give the expression of the unique solution to the problem and we study the general behavior of the minimum Fisher information with respect to the variance. Finally, we give the expression of the probability density with minimum Fisher information among all distributions with finite variance defined on an interval.

2. The differential equation associated with the minimization of Fisher information

In the following, it is convenient to introduce the transformation $f(x) = u(x)^2$ and to work with $u(x)$ instead of $f(x)$. With this notation, the Fisher information becomes

$$I[f] = 4\int_{\mathscr{S}} u'(x)^2\mathrm{d}x, \tag{4}$$

where $u'(x)$ denotes the first order derivative of $u(x)$. The variational problem (3) can be restated as follows:

$$I(\sigma^2) = \inf_{u:u^2\in\mathscr{D}} \left\{4\int_{\mathscr{S}} u'(x)^2\mathrm{d}x : \int_{\mathscr{S}} u(x)^2\mathrm{d}x = 1 \quad \text{and } Var[u^2] = \sigma^2\right\}. \tag{5}$$

The problem above can also be completed by some additional conditions on the boundaries of the domain: for instance, when $\mathscr{S} = [0:+\infty)$ we need $u(+\infty) = 0$ so as to ensure a proper integrable density. At the left endpoint, we shall ensure continuity and set $u(0) = 0$ in order to keep the Fisher information finite. Indeed, the Fisher information associated with a distribution is $+\infty$ when the distribution is not continuous on the domain $\mathscr{S}$. Similarly, when the support $\mathscr{S}$ is restricted to $\mathscr{S} = [-1,1]$, and the density is set to 0 outside of this domain, we shall take $u(1) = u(-1) = 0$ in order to ensure continuity at these boundary points.

The Lagrangian functional associated with the problem (5) is

$$L(u;\alpha,\beta) = \int_{\mathscr{S}} u'(x)^2 \mathrm{d}x + \alpha\left(\int_{\mathscr{S}} u(x)^2 \mathrm{d}x - 1\right) + \beta\left(\int_{\mathscr{S}} (x-\mu)^2 u(x)^2 \mathrm{d}x - \sigma^2\right), \tag{6}$$

where α and β are the Lagrange parameters associated with the normalization and variance constraints respectively. The mean is denoted by μ. The minimum of the Lagrangian functional is obtained by standard calculus of variations [13] which asserts that a minimizer of (6) is necessary a solution of the Euler–Lagrange equation

$$-\frac{d}{dx}\frac{\partial L}{\partial u'} + \frac{\partial L}{\partial u} = 0 \tag{7}$$

which leads to

$$u''(x) - (\alpha + \beta(x-\mu)^2)u(x) = 0. \tag{8}$$

Of course, with the simple change of variable $z = x - \mu$, the differential equation reduces to

$$u''(z) - (\alpha + \beta z^2)u(z) = 0 \tag{9}$$

which is a parabolic differential equation. Interestingly, the minimum Fisher information can be written in terms of the constraints and of the Lagrange parameters associated to these constraints.

Proposition 1. *The minimum Fisher information in (5) can be expressed as*

$$-\frac{1}{4}I(\sigma^2) = \alpha + \beta\sigma^2 \tag{10}$$

with $u(a) = u(b) = 0$ and where a and b denote the left and right endpoints of the support $\mathscr{S}$.

Note that the Lagrange parameters are (complicated) functions of the constraints, so that the right hand side of (10) is not an affine function in σ^2.

Proof. By integration by parts,

$$-\frac{1}{4}I(\sigma^2) = -\int_{\mathscr{S}} u'(x)^2 \mathrm{d}x = -[u(x)u'(x)]_a^b + \int_a^b u(x)u''(x)\mathrm{d}x. \tag{11}$$

Using the boundary conditions and the differential Eq. (8), we then obtain

$$-\frac{1}{4}I(\sigma^2) = \int_a^b (\alpha + \beta(x-\mu)^2)u^2(x)\mathrm{d}x. \tag{12}$$

which reduces to (10) taking into account the values of the constraints. □

3. Solutions to the differential equation

Let us consider the parabolic differential Eq. (9). Using the change of variable $x = \sqrt{2\sqrt{\beta}}z$, together with the substitution $d(x) = u\left((2\sqrt{\beta})^{-\frac{1}{2}}x\right)$, the differential equation becomes

$$d''(x) + \left(-\frac{\alpha}{2\sqrt{\beta}} - \frac{1}{4}x^2\right)d(x) = 0 \tag{13}$$

which is the Weber differential equation, whose standard form is

$$d''(x) + \left(v + \frac{1}{2} - \frac{1}{4}x^2\right)d(x) = 0. \tag{14}$$

Here we simply have $v = -a/2\sqrt{(b)} - 1/2$. The solutions of the Weber equation can be expressed as a linear combination of the parabolic cylinder function $D_v(x)$ and $D_{-v-1}(ix)$ [32]:

$$d(x) = c_1 D_v(x) + c_2 D_{-v-1}(ix) \tag{15}$$

But the Weber equation above can also be converted into the Whittaker equation using the substitution $d(x) = \frac{1}{\sqrt{x}}w\left(\frac{x^2}{2}\right)$ and $z = x^2/2$, which leads to

$$w''(z) + \left[\frac{3}{16z^2} + \frac{v+1/2}{2z} - \frac{1}{4}\right]w(z) \tag{16}$$

that has the form of the Whittaker differential equation

$$w''(z) + \left[\frac{1/4-\mu^2}{z^2} + \frac{\lambda}{z} - \frac{1}{4}\right]w(z) = 0 \tag{17}$$

with $\lambda = \nu/2 + 1/4$ and $\mu = 1/4$. But the Whittaker differential equation can also be obtained directly from the initial differential Eq. (9) with the substitution $u(z) = \frac{1}{\sqrt{z}} w\left(\frac{\xi z^2}{2}\right)$ and with $x = \xi z^2/2$. Then, one readily obtains

$$w''(x) + \left[\frac{3}{16x^2} - \frac{\alpha}{2\xi x} - \frac{\beta}{\xi^2}\right] w(x) = 0, \tag{18}$$

which reduces to the Whittaker differential Eq. (17) with $\xi = 2\sqrt{\beta}, \lambda = -\frac{\alpha}{4\sqrt{\beta}}$, and $\mu = \frac{1}{4}$. The general solutions of the Whittaker differential equation can be expressed in terms of the two linearly independent Whittaker functions $M_{\lambda,\mu}(z)$ and $M_{\lambda,-\mu}(z)$, or with the help of the Whittaker $W_{\lambda,\mu}(z)$ function defined by

$$W_{\lambda,\mu}(z) = \frac{\Gamma(-2\mu)}{\Gamma\left(\frac{1}{2} - \mu - \lambda\right)} M_{\lambda,\mu}(z) + \frac{\Gamma(2\mu)}{\Gamma\left(\frac{1}{2} + \mu + \lambda\right)} M_{\lambda,-\mu}(z) \tag{19}$$

for $2m \notin \mathbb{N}$. In these expressions, the Whittaker $M_{\lambda,\mu}(z)$ function can be expressed as a simple function of the confluent hypergeometric function, or Kummer function, $M\left(\frac{1}{2} + \lambda - \mu, 1 + 2\lambda, z\right)$ according to [1, Eq. 13.1.32]

$$M_{\lambda,\mu}(z) = z^{\lambda+\frac{1}{2}} e^{-\frac{z}{2}} M\left(\frac{1}{2} + \lambda - \mu, 1 + 2\lambda, z\right). \tag{20}$$

Since the Kummer function has an exact series representation, we also have

$$M_{\lambda,\mu}(z) = z^{\mu+\frac{1}{2}} e^{-\frac{z}{2}} \sum \frac{\left(\frac{1}{2} + \lambda - \mu\right)_n}{n!(2\lambda+1)_n} z^n, \tag{21}$$

where $()_n$ denotes the Pochhammer symbol. As far as the parabolic cylinder function is concerned, it can also be written as a Whittaker function [32, p. 347]:

$$D_\nu(x) = 2^{\frac{\nu}{2}+\frac{1}{4}} z^{-\frac{1}{2}} W_{\frac{\nu}{2}+\frac{1}{4},-\frac{1}{4}}\left(\frac{1}{2}z^2\right). \tag{22}$$

According to this discussion and the relationships between parabolic cylinder functions, Whittaker M and W functions, we find that the solutions of the differential equation associated with the problem of minimum Fisher information can be expressed as various equivalent linear combinations of the functions

$$D_{-\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}} z\right) \text{ and } D_{\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}} iz\right),$$

or

$$\frac{1}{\sqrt{z}} M_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}z^2\right), \quad \frac{1}{\sqrt{z}} M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}\left(\sqrt{\beta}z^2\right), \quad \text{and } \frac{1}{\sqrt{z}} W_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}z^2\right)$$

Some equivalent expressions of the solutions are given below

$$u(z) = c_1 D_{-\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}} z\right) + c_2 D_{\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}} iz\right) \tag{23}$$

$$= c_1' \frac{1}{\sqrt{z}} M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}\left(\sqrt{\beta}z^2\right) + c_2' \frac{1}{\sqrt{z}} M_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}z^2\right) \tag{24}$$

$$= c_1'' \frac{1}{\sqrt{z}} M_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}z^2\right) + c_2'' \frac{1}{\sqrt{z}} W_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}z^2\right). \tag{25}$$

In these formulas, the values of the constants in the linear combinations will be determined according to auxiliary constraints, e.g. the boundary values. We now turn to the characterization of solutions on $\mathbb{R}^+$ and then on an interval.

4. Solutions defined on the positive real line

We consider the expression of the solution in terms of the parabolic cylinder functions, as given in (23). A first point is to check that the solution is bounded on $\mathbb{R}^+$. Since the Weber equation has only one irregular singularity at $z = +\infty$, it is sufficient to examine the behavior of the solution for $z \to +\infty$. The asymptotic expansion of the Weber function $D_\nu(z)$, with $\arg z \leqslant 3\pi/4$ is [32, p. 347],[1, Eq. 19.8.1]:

$$D_\nu(z) \sim e^{-\frac{1}{4}z^2} z^\nu \left(1 - \frac{\nu(\nu-1)}{2z^2} + \frac{\nu(\nu-1)(\nu-2)(\nu-3)}{24z^4} - \cdots\right). \tag{26}$$

For $z \to +\infty$, we see at once from the asymptotic expansion that $D_\nu(z) \to 0$. Replacing ν by $-\nu - 1$ and z by iz, we can also observe that $D_{-\nu-1}(iz)$ tends to infinity when $z \to +\infty$. Therefore, since the solution must correspond to a proper integrable density, the second parabolic cylinder function in (23) must be discarded, with $c_2 = 0$, and the solution becomes

$$u_{\alpha,\beta}(z) = \frac{D_{-\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}}z\right)}{\left(\int_0^{+\infty}\left(D_{-\frac{\alpha}{2\sqrt{\beta}}-\frac{1}{2}}\left(\left(2\sqrt{\beta}\right)^{\frac{1}{2}}z\right)\right)^2 \mathrm{d}z\right)^{\frac{1}{2}}} \tag{27}$$

where the denominator has been introduced in order to ensure the normalization of the probability density $f(z) = u(z)^2$.

As a consequence of the simple form obtained above, it is possible to characterize the general behavior of the minimum Fisher information as a function of the variance. This is given by the following

Proposition 2. *The Fisher information of the minimum information probability density with positive support and fixed variance σ^2, corresponding to the solution (27), verifies*

$$I_\nu(\sigma^2) = \frac{K_\nu}{\sigma^2}, \tag{28}$$

where $\nu = -\alpha/2\sqrt{\beta} - 1/2$ and K_ν is a constant.

Proof. Let $\nu = -\alpha/2\sqrt{\beta} - 1/2$ and $\xi = 2\sqrt{\beta}$, and consider ν fixed. Then, ξ acts as a scaling factor: more precisely, denoting $V_{\nu,\xi}$ the variance associated with the solution with parameters ν and ξ, we readily have $V_{\nu,\xi} = \frac{1}{\xi}V_{\nu,1}$ and $I_{\nu,\xi} = \xi I_{\nu,1}$, where $I_{\nu,\xi}$ is the Fisher information. Therefore, the product $I_{\nu,\xi}V_{\nu,\xi} = I_{\nu,1}V_{\nu,1}$ does not depend on ξ and

$$I_{\nu,\xi} = \frac{I_{\nu,1}V_{\nu,1}}{V_{\nu,\xi}}.$$

Moreover, from the equality $V_{\nu,\xi} = \frac{1}{\xi}V_{\nu,1}$ and the fact that $V_{\nu,1} > 0$, we deduce that the function $\xi \mapsto V_{\nu,\xi}$ maps $\mathbb{R}^+$ to $\mathbb{R}^+$ so that it is always possible to find a value ξ such that $V_{\nu,\xi} = \sigma^2$; thus, the Fisher information $I_\nu(\sigma^2)$ of the probability density associated with the solution $u(z)$ in (27) follows the Eq. (28), with $K_\nu = I_{\nu,1}V_{\nu,1}$. □

Let ν^* be the value that minimizes $I_\nu(\sigma^2)$. For that value and any distribution f on $\mathbb{R}^+$ with same variance σ^2, we always have

$$I[f] \geqslant I_{\nu^*}(\sigma^2) = \frac{K_{\nu^*}}{\sigma^2}, \tag{29}$$

which refines the Cramér-Rao inequality. We will check below that K_{ν^*} is of course bigger than one.

Actually, we know that the Fisher information is infinite in the case of a non differentiable density. It is thus important here to ensure continuity and differentiability at the origin. In order to ensure that the Fisher information of the probability distribution $f(x) = u^2(x)$ remains finite, we need to impose $u(0) = 0$. This implies a condition on ν so that $D_\nu(0) = 0$. It is easy to check that

$$D_\nu(0) = \frac{\sqrt{\pi}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{1}{2} - \frac{1}{2}\nu\right)}. \tag{30}$$

Therefore, $D_\nu(0) = 0$ if and only if $\Gamma\left(\frac{1}{2} - \frac{1}{2}\nu\right) \to +\infty$, that is if and only if ν is an odd positive integer. In such a case, the Weber functions can also be expressed in terms of Hermite polynomials $H_n(x)$:

$$D_n(x) = 2^{-\frac{n}{2}}e^{-\frac{x^2}{4}}H_n(x). \tag{31}$$

Then, the determination of the optimum value ν^* of ν such that $I_\nu(\sigma^2)$ is minimum amounts to minimize K_ν with ν integer. This leads to the following result.

Proposition 3. *For a given variance σ^2, the distribution on $\mathbb{R}^+$ which minimizes the Fisher information is obtained for $\nu^* = 1$, and is*

$$f_\xi(x) = \sqrt{\frac{2}{\pi}}\xi^{\frac{3}{2}}x^2 \exp\left(-\frac{\xi x^2}{2}\right), \tag{32}$$

which is the chi-distribution with three degrees of freedom, and where the parameter ξ is given by $\xi = \left(3 - \frac{8}{\pi}\right)/\sigma^2$. Its Fisher information, according to (45), is $I_{1,\xi} = 3\xi$. Then the minimum Fisher-variance product is $K_1 = 9 - 24/\pi \approx 1.3606$.

Proof. In the present case, it is possible to obtain closed-form formulas for the variance and information associated with (27), even for non integer values of ν.

The following relationships, which can be derived from integral representations of parabolic cylinder functions, see [31, 6.2] are useful:

$$\frac{d}{dx}D_\nu(x) = (-1/2)xD_\nu(x) + \nu D_{\nu-1}(x) \tag{33}$$

$$= (1/2)xD_\nu(x) - D_{\nu+1}(x) \tag{34}$$

Adding the two equalities and taking the square, we have

$$4\left(\frac{d}{dx}D_\nu(x)\right)^2 = \nu^2 D_{\nu-1}(x)^2 - 2\nu D_{\nu-1}(x)D_{\nu+1}(x) + D_{\nu+1}(x)^2. \tag{35}$$

Subtracting the Eq. (33) and (34), we obtain

$$xD_\nu(x) = D_{\nu+1}(x) - \nu D_{\nu-1}(x), \tag{36}$$

from which we deduce

$$xD_\nu(x)^2 = D_{\nu+1}(x)D_\nu(x) - \nu D_{\nu-1}(x)D_\nu(x), \tag{37}$$

$$x^2 D_\nu(x)^2 = D_{\nu+1}(x)^2 - 2\nu D_{\nu+1}(x)D_{\nu-1}(x) + \nu^2 D_{\nu-1}(x)^2, \tag{38}$$

The second ingredient of the calculation are the formulas [14, Eqs. 7.711.2 and 7.711.3], from which we define the functional $S(\mu,\nu)$:

$$S(\mu,\nu) = \int_0^{+\infty} D_\mu(x)D_\nu(x)\mathrm{d}x = \frac{\pi 2^{\mu+\nu+1}}{\mu-\nu}\left[\frac{1}{\Gamma\left(\frac{1}{2}-\frac{1}{2}\mu\right)\Gamma\left(-\frac{1}{2}\nu\right)} - \frac{1}{\Gamma\left(\frac{1}{2}-\frac{1}{2}\nu\right)\Gamma\left(-\frac{1}{2}\mu\right)}\right] \tag{39}$$

for $\mu \neq \nu$, and with

$$S(\nu,\nu) = \int_0^{+\infty} D_\nu(x)^2\mathrm{d}x = \pi^{\frac{1}{2}}2^{-\frac{3}{2}}\frac{\Psi\left(\frac{1}{2}-\frac{1}{2}\nu\right) - \Psi\left(-\frac{1}{2}\nu\right))}{\Gamma(-\nu)} \tag{40}$$

where $\Psi(x)$ is the digamma function. When μ or ν are integers, these formulas reduce to

$$S(2p,2p+k) = -\frac{\pi 2^{2p+\frac{1}{2}k+\frac{1}{2}}}{k}\frac{1}{\Gamma\left(\frac{1}{2}-p\right)\Gamma\left(-p-\frac{1}{2}k\right)} \tag{41}$$

$$S(2p+1,2p+1+k) = \frac{\pi 2^{2p+\frac{1}{2}k+\frac{3}{2}}}{k}\frac{1}{\Gamma\left(-\frac{1}{2}-p\right)\Gamma\left(-p-\frac{1}{2}k\right)} \tag{42}$$

$$\text{and } S(m,m) = (2\pi)^{\frac{1}{2}}\frac{m!}{2} \tag{43}$$

So doing, using the expression of the solution (27), the definition of the Fisher information (4), equality (35) and the definitions (39) and (40), we obtain the expression

$$I_{\nu,\xi} = \xi(\nu^2 S(\nu-1,\nu-1) - 2\nu S(\nu-1,\nu+1) + S(\nu+1,\nu+1))/S(\nu,\nu), \tag{44}$$

which reduces to the very simple expression

$$I_{m,\xi} = (2m+1)\xi \tag{45}$$

in the integer case. Let us mention that a similar expression is reported in [36] for the case of Hermite functions, up to a factor 2 which is due to a different definition of Hermite functions. In the same way, the integration of equalities (37) and (38) gives the first and second order moment, so the variance is

$$V_{\nu,\xi} = \frac{1}{\xi}(S(\nu+1,\nu+1) - 2\nu S(\nu+1,\nu-1) + \nu^2 S(\nu-1,\nu-1))/S(\nu,\nu) - ((S(\nu+1,\nu) - \nu S(\nu-1,\nu))/S(\nu,\nu))^2 \tag{46}$$

Figs. 1 and 2 present the evolution of the variance and information, as computed in (44) and (46). Fig. 3 gives the information-variance product K_ν, as a function of ν. Actually, we know that the true Fisher information is only finite for positive odd values of ν (otherwise the density is discontinuous at the origin and the Fisher information is infinite).

Considering Fig. 3, we read that the positive odd integer which minimizes the product $I_{\nu,1}\cdot V_{\nu,1}$ is $\nu^* = 1$. Accordingly, we obtain that the solution $f(x) = u(x)^2$ is (32). The values of its variance and Fisher information follow by direct computation. □

Note that disregarding the differentiability requirement at the origin would lead to select the parameter $\nu = 0.1065$, corresponding to a variance $\sigma^2 = 0.38661$ and a "Fisher information" $I = 0.91886$. So doing one would obtain a product $I\cdot\sigma^2 = 0.35524$, which would break the Cramer-Rao inequality. From the estimation theory point of view, it is clear that if the density has a bounded support and a discontinuity at the left endpoint of this support, then the variance of the estimate of the location parameter will be asymptotically zero, which corresponds to an infinite Fisher information in the Cramér-Rao bound. Clearly, the estimator defined as the minimum of the experimental data converges to the value of the left endpoint of the support and, in turn, provides an estimate of the value of the location parameter with asymptotically zero-variance.

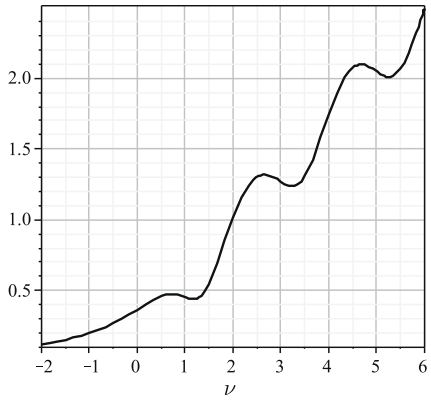


Fig. 1. Evolution of the variance $V_{\nu,1}$ in (44) with respect to ν.

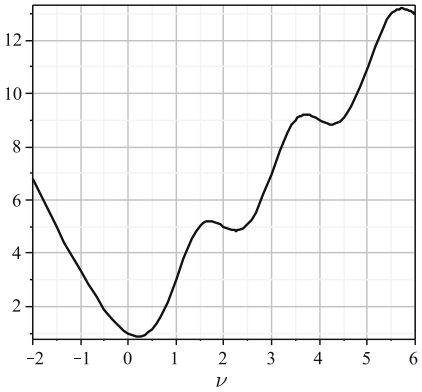


Fig. 2. Evolution of the information $I_{\nu,1}$ in (46) with respect to ν. Actually, the true Fisher information is only finite for positive odd values of ν.

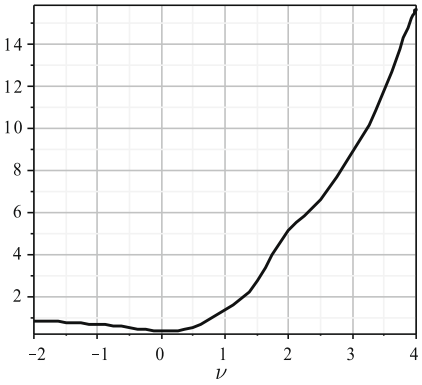


Fig. 3. The information-variance product K_ν as a function of ν.

5. Solutions with compact support

We know that the Fisher information associated with a distribution is invariant by translation of this distribution. Furthermore, a scaling of the distribution $f(x)$ according to $g(x) = \frac{1}{|a|} f\left(\frac{x}{a}\right)$ with $a \neq 0$ yields a scaling of the Fisher information as $I[g] = \frac{1}{|a|^2} I[f]$. Hence, it is possible to restrict our study to any particular interval, and without loss of generality, we choose the interval $[-1, 1]$.

We consider the same problem as before, the minimization of the Fisher information subject to a variance constraint, and add the boundaries conditions $u(1) = u(-1) = 0$, which reads

$$I(\sigma^2) = \inf_{u:u^2\in\mathscr{D}} \left\{ 4\int_{-1}^{1} u'(x)^2 \mathrm{d}x : \int_{-1}^{1} u(x)^2 \mathrm{d}x = 1 \text{ with } Var(u^2) = \sigma^2 \text{ and } u(1) = u(-1) = 0 \right\} \tag{47}$$

The problem is invariant by the symmetry $x \mapsto -x$ since it is clear that if $u(x)$ is solution, then $v(x) = u(-x)$, which has the same variance and Fisher information, is also a solution. In fact, the distribution with minimum Fisher information for a given variance is unique and even. The general proof of this fact is rather involved and given in [27,28]. But since we have here the general expression of solutions to the underlying differential equation, it is not difficult to characterize this optimum solution.

Proposition 4. *The unique solution to the problem (47) is the non-negative function*

$$u(x) = \frac{\frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}\left(\sqrt{\beta}x^2\right)}{\left(\int_{-1}^{1}\left(\frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}\left(\sqrt{\beta}x^2\right)\right)^2 \mathrm{d}x\right)^{\frac{1}{2}}}, \tag{48}$$

where $\sqrt{\beta}$ is the first zero of the function $M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}$.

Proof. We have already established that the general solution of the differential Eq. (9) can be expressed as a linear combination of two Whittaker M functions which are two linearly independent solutions:

$$u(x) = c_1 \frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}\left(\sqrt{\beta}x^2\right) + c_2 \frac{1}{\sqrt{x}} M_{-\frac{\alpha}{4\sqrt{\beta}},\frac{1}{4}}\left(\sqrt{\beta}x^2\right), \tag{49}$$

where c_1 and c_2 shall be chosen such that the boundaries and normalization conditions are satisfied. In this last equation, the first function is even while the second is odd, as it can be observed from the series development (21). Let us note them $S_{\text{even}}(x)$ and $S_{\text{odd}}(x)$, respectively. The boundaries conditions $u(1) = u(-1) = 0$ then imply that

$$c_1 S_{\text{even}}(1) + c_2 S_{\text{odd}}(1) = 0, \tag{50}$$

$$c_1 S_{\text{even}}(1) - c_2 S_{\text{odd}}(1) = 0. \tag{51}$$

The only solution is $c_1 = c_2 = 0$, except if $S_{\text{odd}}(1)$ and $S_{\text{even}}(1)$ are simultaneously equal to zero. But it is easy to check that these two functions have no common zero. Therefore, the general solution can not include both terms. There shall be only one term, either $S_{\text{even}}(x)$ or $S_{\text{odd}}(x)$, in the general expression of the solution with its parameters adjusted so as to ensure the boundaries constraints.

We deduce that, since $u(x)$ is either odd or even, the associated density $f(x) = u^2(x)$ is even and has zero mean. In such a case, the set defined by the constraint $\mathrm{Var}[f] = \sigma^2$ becomes a convex set and consequently the solution to the minimization of the Fisher information on this convex set is unique. From the fact both $u(x)$ and $v(x) = u(-x)$ are solution, we deduce that the unique solution is necessary the even one, that is (48).

Finally, if $u(x)$ is solution, so is $w(x) = |u(x)|$ because $w'(x)^2 = u'(x)^2$ and $w(x)^2 = u(x)^2$. By uniqueness of the solution we obtain that $u(x) \geqslant 0$ for $x \in (-1, 1)$. Therefore the condition $u(1) = 0$, the fact that all zeros of $M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}(x)$ are zero-crossings, and the non-negativity requirement, yield that the point $\sqrt{\beta}$ is the first zero of $M_{-\frac{\alpha}{4\sqrt{\beta}},-\frac{1}{4}}$. □

In Fig. 4, we give some examples of solutions for several values $\sigma^2 \in [0, 1]$ of the variance. We see that for low variances the solution is unimodal while it is bimodal for higher variances. The Fisher information is large for low and high variances and much lower for intermediate variances. In fact, the minimum Fisher information is a convex function of the variance, cf [28, Theorem 6.1].

Proposition 5. *The minimum Fisher information $I(\sigma^2)$ is a strictly convex function of σ^2.*

Proof. Let $u(x)^2$ and $v(x)^2$ be two distributions with minimum informations $I(\sigma_u^2)$ and $I(\sigma_v^2)$ respectively. Since the Fisher information is strictly convex, with $\epsilon \in (0, 1)$, we have

$$I[\epsilon u^2 + (1-\epsilon)v^2] < \epsilon I[u^2] + (1-\epsilon) I[v^2] = \epsilon I(\sigma_u^2) + (1-\epsilon) I(\sigma_v^2) \tag{52}$$

Since the distributions u^2 and v^2 have zero mean, we have

$$Var[\epsilon u^2 + (1-\epsilon)v^2] = \epsilon\sigma_u^2 + (1-\epsilon)\sigma_v^2.$$

Then, there exists a distribution with the same variance and minimum Fisher information such that

$$I[\epsilon u^2 + (1-\epsilon)v^2] \geqslant I(\epsilon\sigma_u^2 + (1-\epsilon)\sigma_v^2). \tag{53}$$

Finally, combination of (52) and (53) yields

$$I(\epsilon\sigma_u^2 + (1-\epsilon)\sigma_v^2) < \epsilon I(\sigma_u^2) + (1-\epsilon) I(\sigma_v^2). \qquad \square \tag{54}$$

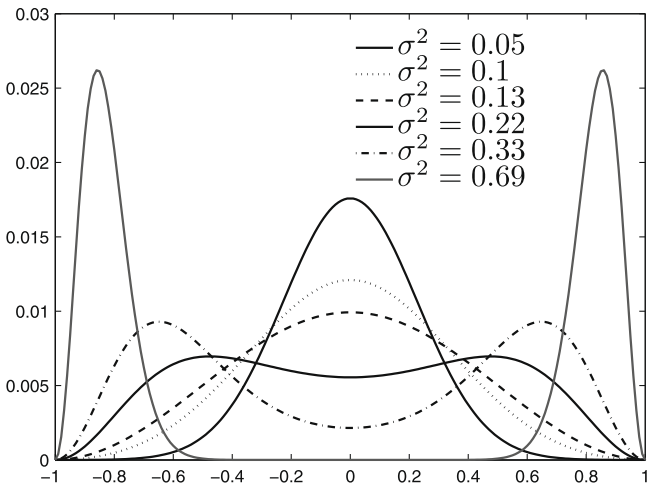


Fig. 4. Minimum Fisher information distributions with compact $(0,1)$ support for several values of variances. For variances lower than $\sigma_*^2 = \frac{1}{3} - \frac{2}{\pi^2} \approx 0.1307$, the distributions are unimodal while they become bimodal when $\sigma^2 > \sigma_*^2$.

In the case of the $\mathbb{R}^+$ support, we obtained explicit expressions of both the variance and Fisher information, and as a result the exact behavior of $I(\sigma^2)$. The present case is more delicate. Indeed, although it is possible to obtain from (21) an exact series expansion of $u^2(x)$, where the coefficients depend of hypergeometric functions, the difficulty here is that we have no analytical expression for the argument of the first zero of the solution, which gives the value of β. Only approximations of this value are available [1, p. 510]. Therefore, we have to resort to a numerical determination of the parameters of the function. Similarly, we cannot give a close form formula for $I(\sigma^2)$. However, we can still characterize its general behavior.

Since $I(\sigma^2)$ is a convex function, it has a unique minimum, say σ_*^2. The following result shows that this minimum is obtained for $\beta = 0$. Furthermore, this value discriminates two regimes for the solutions: in the case $\sigma < \sigma_*$, we have $\beta > 0$ while in the case $\sigma > \sigma_*$ we have $\beta < 0$ and the corresponding solution is the Whittaker M function with imaginary arguments. These facts have already been noticed in [28, Lemmas 5.2 and 5.3], but we provide here an alternate proof.

Proposition 6. *If σ_*^2 is the minimizer of $I(\sigma^2)$, then*

(a) *for* $\sigma < \sigma_*, \quad \beta > 0$
(b) *for* $\sigma > \sigma_*, \quad \beta < 0$
(c) *and finally for* $\sigma = \sigma_*, \quad \beta = 0.$

Proof. Let $u(x)^2$ and $v(x)^2$ be two distributions with minimum informations $I(\sigma_u^2)$ and $I(\sigma_v^2)$ respectively, and let us define $g_\epsilon(x) = \epsilon u^2(x) + (1-\epsilon)v^2(x)$. By Proposition 1, the Fisher information $I(\sigma^2)$ can be expressed as $I(\sigma^2) = -4(\alpha_v + \beta_v \sigma_v^2)$. We use the expansion proved in [27, Satz 7.2, p. 90]:

$$I[g_\epsilon] = I[v^2] - 4\beta_v \epsilon(\sigma_u^2 - \sigma_v^2) + o(\epsilon^2). \tag{55}$$

Strict convexity gives

$$I[g_\epsilon] < \epsilon I[u^2] + (1-\epsilon) I[v^2] = \epsilon I(\sigma_u^2) + (1-\epsilon) I(\sigma_v^2). \tag{56}$$

Let us now take $\sigma_u = \sigma_*$. Since $I(\sigma_*^2)$ is the minimum Fisher information, we have the majorization

$$\epsilon I(\sigma_*^2) + (1-\epsilon) I(\sigma_v^2) \leqslant I(\sigma_v^2), \tag{57}$$

and therefore $I[g_\epsilon] < I(\sigma_v^2)$. As a consequence, from (55) and the previous inequality, we obtain

$$\beta_v \epsilon(\sigma_*^2 - \sigma_v^2) > 0, \tag{58}$$

which gives cases (a) and (b) in the Proposition. The case (c), $\beta = 0$, follows by continuity. □

Hence, the solution which realizes the minimum $I(\sigma_*^2)$ of the Fisher information corresponds to $\beta = 0$. This means that this solution satisfies the differential equation $u''(x) = \alpha u(x)$, with $u(1) = u(-1) = 0$. This problem has as solutions $u(x) = \cos(k\pi x/2)$, with $\alpha = -k^2\pi^2$ and k integer. Since we know that the solution of the Fisher minimization is non-negative, there is only one possibility and

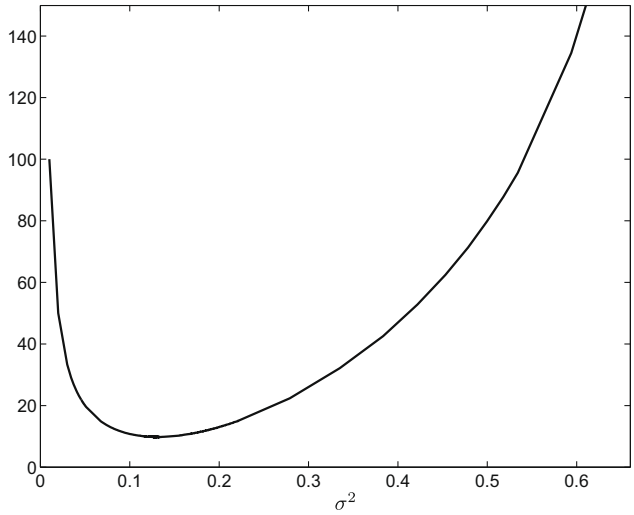


Fig. 5. Evolution of the minimum Fisher information $I(\sigma^2)$ in the case of a distribution with bounded support $(-1,1)$.

Proposition 7. *The probability density function $f(x) = u(x)^2$ defined on $[-1,1]$ with minimum Fisher information is*

$$f(x) = \cos^2(\pi x/2) \tag{59}$$

$$\text{with } \sigma_*^2 = \frac{1}{3} - \frac{2}{\pi^2} \text{ and } I(\sigma_*^2) = \pi^2. \tag{60}$$

Fig. 5 reports the behavior of the Fisher information $I(\sigma^2)$. The Fisher information tends to infinity for $\sigma \to 0$ and $\sigma \to +\infty$ and its minimum is attained for $\sigma^2 = \sigma_*^2$. For $\sigma \to 0$, we obtain by [1, Eqs. 13.1.32 and 13.5.5] that $u(x) \propto \exp\left(-\beta^{\frac{1}{2}}x^2/2\right)$, that is $u(x)$ converges to a Gaussian distribution with variance $\beta^{-\frac{1}{2}}$. Hence, for small values of the variance, the solution has the form of a Gaussian distribution concentrated on the origin, and, as a normal distribution, its Fisher information decreases as $1/\sigma^2$. When σ^2 increases towards 1, the two modes become more and more pronounced and probability accumulates on ± 1, and the probability density function tends to two mass functions, as shown in Fig. 4, and the Fisher information tends to infinity.

6. Conclusion

In this paper, we have solved the problem of minimizing the Fisher information on restricted supports with a fixed variance. The problem has been stated under the form of a general second order linear differential equation. We have shown that the general form of the solutions involves Whittaker functions. We have derived the explicit expressions of the solutions on $\mathbb{R}^+$ and on an interval. We have first studied the set of solutions on $\mathbb{R}^+$ and shown that the distribution with minimum Fisher information is a scaled chi distribution. On the interval $[-1,+1]$, we have characterized the solutions, investigated their behavior, and shown that the distribution with minimum Fisher information is a squared cosine function.

Future work will consider the extension of these results in the multivariate case. We also intend to investigate some interesting generalized versions of Fisher information [18,4] as suggested by one of the referees of this paper.

Acknowledgements

The authors greatly acknowledge Elke Uhrmann-Klingen for friendly correspondence and for communicating us her manuscript [27]. C.V. acknowledges interesting conversations with R. Frieden and B. Soffer during the 3rd International Workshop on Fisher Information held in Tucson, April 2008.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables, ninth Dover printing, tenth GPO printing ed., Dover Publications, 1964.
[2] M.M. Ali, S. Nadarajah, Information matrices for normal and Laplace mixtures, Information Sciences 177 (3) (2007) 947–955.
[3] M. Balasco, V. Lapenna, M. Lovallo, G. Romano, A. Siniscalchi, L. Telesca, Fisher information measure analysis of earth's apparent resistivity, International Journal of Nonlinear Science 5 (3) (2008) 230–236.
[4] D. Boekee, An extension of the Fisher information measure, in: I. Csiszár, P. Elias (eds.), Topics in Information Theory Proc. 2nd Colloquium on Information Theory; Keszthely, Hungary, 25–29 August 1975, vol. 16, János Bolyai Mathematical Society and North-Holland, Keszthely, Hungary, 1977.

[5] J.M. Borwein, A.S. Lewis, M.N. Limber, D. Noll, Maximum entropy reconstruction using derivative information part 2: computational results, Numerische Mathematik 69 (3) (1995) 243–256.
[6] J.M. Borwein, A.S. Lewis, D. Noll, Maximum entropy reconstruction using derivative information, part 1: Fisher information and convex duality, Mathematics of Operations Research 21 (1996) 442–468.
[7] M. Cohen, The Fisher information and convexity, IEEE Transactions on Information Theory 14 (4) (1968) 591–592.
[8] A. Dasgupta, Stein's identity, Fisher information, and projection pursuit: a triangulation, Journal of Statistical Planning and Inference 137 (11) (2007) 3394–3409.
[9] S.A. Frank, Natural selection maximizes Fisher information, Journal of Evolutionary Biology 22 (2) (2009) 231–244.
[10] B.R. Frieden, Fisher information, disorder, and the equilibrium distributions of physics, Physical Review A 41 (8) (1990) 4265–4276.
[11] B.R. Frieden, Science from Fisher Information: A Unification, Cambridge University Press, 2004.
[12] B.R. Frieden, P.M. Binder, Physics from Fisher information: a unification, American Journal of Physics 68 (2000) 1064.
[13] I.M. Gelfand, S.V. Fomin, Calculus of Variations, Dover Publications, 2000.
[14] I. Gradshteyn, I. Ryzhik, Table of Integrals, Series and Products, Academic P, 1966.
[15] P.J. Huber, Robust statistics: a review, The Annals of Mathematical Statistics 43 (3) (1972) 1041–1067.
[16] M. Lovallo, F. Marchese, N. Pergola, L. Telesca, Fisher information measure of temporal fluctuations in satellite advanced very high resolution radiometer (AVHRR) thermal signals recorded in the volcanic area of Etna (Italy), Communications in Nonlinear Science and Numerical Simulation 14 (1) (2009) 174–181.
[17] M.T. Martin, F. Pennini, A. Plastino, Fisher's information and the analysis of complex signals, Physics Letters A 256 (2-3) (1999) 173–180.
[18] A. Mathai, Basic Concepts in Information Theory and Statistics: Axiomatic Foundations and Applications, Wiley, New York, 1975.
[19] M. Montenegro, M.R. Casals, M.A. Lubiano, M. Ángeles Gil, Two-sample hypothesis tests of means of a fuzzy random variable, Information Sciences 133 (1-2) (2001) 89–100.
[20] A. Nagy, Fisher information in density functional theory, The Journal of Chemical Physics 119 (18) (2003) 9401–9405.
[21] S. Nordebo, M. Gustafsson, B. Nilsson, Fisher information analysis for two-dimensional microwave tomography, Inverse Problems 23 (3) (2007) 859–877.
[22] E. Romera, J.S. Dehesa, Weizsäcker energy of many-electron systems, Physical Review A 50 (1) (1994) 256–266.
[23] K.D. Sen, J. Antolín, J.C. Angulo, Fisher–Shannon analysis of ionization processes and isoelectronic series, Physical Review A 76 (3) (2007) 32502.
[24] J. Sánchez-Ruiz, J.S. Dehesa, Fisher information of orthogonal hypergeometric polynomials, Journal of Computational and Applied Mathematics 182 (1) (2005) 150–164.
[25] L. Telesca, V. Lapenna, M. Lovallo, Fisher information analysis of earthquake-related geoelectrical signals, Natural Hazards and Earth System Sciences 5 (4) (2005) 561–564.
[26] C. Tsairidis, K. Ferentinos, T. Papaioannou, Information and random censoring, Information Sciences 92 (1–4) (1996) 159–174.
[27] E. Uhrmann-Klingen, Fisher-Minimale Dichten auf Kompacten intervallen (in German), Ph.D. thesis, Essen, 1992.
[28] E. Uhrmann Klingen, Minimal Fisher information distributions with compact-supports, Sankhya: The Indian Journal of Statistics 57 (3) (1995) 360–374.
[29] C. Vignat, J.F. Bercher, Analysis of signals in the Fisher–Shannon information plane, Physics Letters A 312 (1-2) (2003) 27–33.
[30] C. Vignat, J.-F. Bercher, On Fisher information inequalities and score functions in non-invertible linear systems, Journal of Inequalities in Pure and Applied Mathematics 4 (4) (2003) 71.
[31] Z.X. Wang, D.R. Guo, Special Functions, World Scientific Pub. Co. Inc., 1989.
[32] E.T. Whittaker, G.N. Watson, A Course of Modern Analysis, 4th ed., Cambridge University Press, 1927.
[33] E.K.H. Wu, Distributions minimizing Fisher information for scale in [var epsilon]-contamination neighbourhoods, Statistics & Probability Letters 14 (5) (1992) 373–383.
[34] E.K.H. Wu, P.S. Chan, Distributions minimizing Fisher information for location in Kolmogorov neighbourhoods, Annals of the Institute of Statistical Mathematics 49 (3) (1997) 541–554.
[35] B. Xu, Q. Chen, Z. Wu, Z. Wang, Analysis and approximation of performance bound for two-observer bearings-only tracking, Information Sciences 178 (8) (2008) 2059–2078.
[36] R.J. Yáñez, P. Sánchez-Moreno, A. Zarzo, J.S. Dehesa, Fisher information of special functions and second-order differential equations, Journal of Mathematical Physics 49 (2008) 082104.
[37] V. Zivojnovic, Minimum Fisher information of moment-constrained distributions with application to robust blind identification, Signal Processing 65 (2) (1998) 297–313.
[38] V. Zivojnovic, D. Noll, Minimum Fisher information spectral analysis, in: ICASSP-97, IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, 1997.

8.9 C. Berland, I. Hibon, J.-F. Bercher et al. « New transmitter architecture for nonconstant envelope modulation », *IEEE Trans. on Circuits and Systems*, vol. 53, Issue 1, pp 13 - 17, 2006.

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS, VOL. 53, NO. 1, JANUARY 2006 13

A Transmitter Architecture for Nonconstant Envelope Modulation

C. Berland, *Member, IEEE*, I. Hibon, J. F. Bercher, M. Villegas, D. Belot, D. Pache, and V. Le Goascoz

***Abstract*—This paper presents a transmitter architecture which is less sensitive to power amplifier AM/AM and AM/PM conversions. Indeed, most radio communication standards employ nonconstant envelope modulation such as in the 802.11 series. Owing to power amplifiers' nonlinearities, distortions on the modulations degrade the overall transmitter performance. As a consequence, linearization methods, such as envelope elimination and restoration, have to be employed. The architecture relies on the transformation of a nonconstant envelope modulation to a constant or switched one. This signal coding enables the efficient amplification of the signal without undergoing power amplifier distortions. The envelope variation is restored by bandpass filtering just after the power amplifier.**

***Index Terms*—Linearization techniques, pulsewidth modulation (PWM), sigma–delta ($\Sigma\Delta$) modulation, transceivers.**

I. Introduction

ENVELOPE variation in digital modulation increases transmitter design complexity. AM/AM and AM/PM conversion effects of the power amplifier (PA) degrade error vector magnitude (EVM) and the output spectrum of the transmitted signal. Therefore, linearization methods, one of which is envelope elimination and restoration (EER), have to be implemented. EER [1] is an analog solution that relies upon the separation of the envelope and phase parts of the modulated signal. The envelope signal is restored through its injection via the PA supply voltage. Pulsewidth modulation (PWM) of the envelope and amplification via a Class-D amplifier achieve this operation, as demonstrated in Fig. 1. The dc–dc converter implementation is complex because of the envelope variation and bandwidth it has to handle.

Advances in technologies, together with the significant improvement in the frequency capabilities of heterojunction bipolar and CMOS, now enable the design of high frequency switched circuits, e.g a 8-GHz sigma–delta ($\Sigma\Delta$) [2], or a 16-GHz comparator [3].

These advances allow the classical EER architecture to be adapted using such high-frequency (HF) circuits to produce a new solution which is more efficient and has beneficial characteristics.

Manuscript received September 14, 2004; revised February 16 2005. This paper was recommended by Associate Editor F. C. M. Lau.

C. Berland, J. F. Bercher, and M. Villegas are with ESYCOM-ESIEE laboratoire système de communication, 93162 Noisy-Le-Grand Cedex, France (e-mail: c.berland@esiee.fr).

I. Hibon, D. Belot, D. Pache, and V. Le Goascoz are with ST Microelectronics, 38926 Crolles Cedex, France (e-mail: isabelle.hibon@st.com).

Digital Object Identifier 10.1109/TCSII.2005.854594

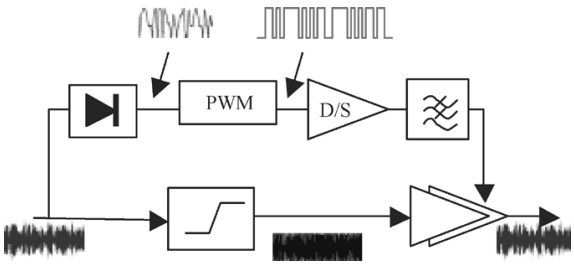


Fig. 1.PWM and restoration of the envelope.

The classical EER architecture can be modified in such a way that the restoration occurs before the power amplification by a simple multiplication between the constant envelope phase signal and that of the PWM. The envelope variation is recovered after the power amplifier by bandpass filtering. Such an architecture using a "0/1 $\Sigma\Delta$" was proposed recently in [4]. Another development of the original EER is given in [5] and presented here. In order to avoid degradations related to the envelope detector and to the limiter, we choose to generate digitally both phase and envelope signals, as indicated in Fig. 2.

A significant improvement is obtained using a "$\pm a$ $\Sigma\Delta$" which results in a constant envelope signal at the PA input as shown in Fig. 3. Consequently, the choice of PA is not reduced to the single switched PA class. This solution also simplifies the implementation of the PA dc–dc converter.

This paper presents performances of this new transmitter architecture applied to a 16 QAM and an orthogonal frequency-division multiplexing (OFDM) modulation. The discussion will be mainly focused on the PWM generation (classical in Section III and using a 1-bit "$\pm a$ $\Sigma\Delta$" in Section IV) and its impact on characteristics of other transmitter components. In both cases, we present the architecture and related properties, its tuning, and its performances. Finally, the conclusion highlights the benefits of this approach.

II. Preliminary—Introduction to OFDM Modulation

Due to its properties in terms of bandwidth and envelope magnitude variation, the OFDM modulation of hiperlan2 standard is chosen here to validate the transmitter architecture. The transmission frequency is about 5 GHz, with a 16-QAM symbol time duration of 50 ns. For comparison purposes, the architecture is also studied in the case of a simple 16-QAM modulation with the same symbol time duration.

The OFDM principle consists in the multiplication of subcarriers with QAM symbols. The time-domain signal which is

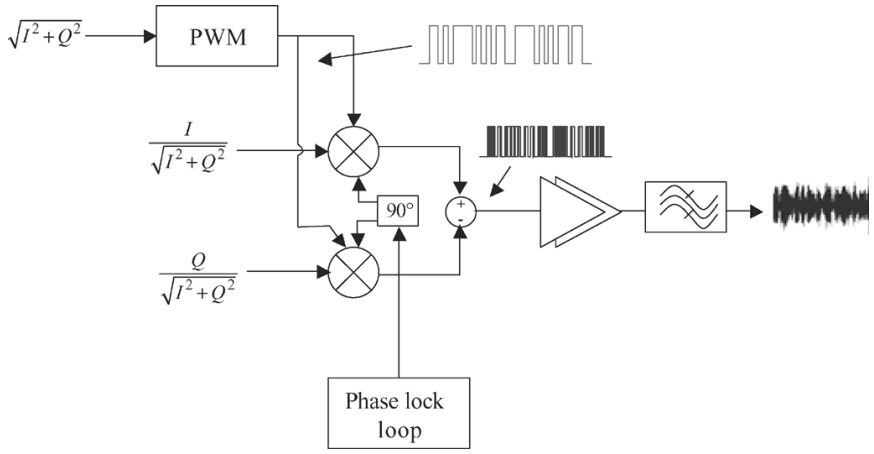


Fig. 2. Principle of the new transmitter architecture.

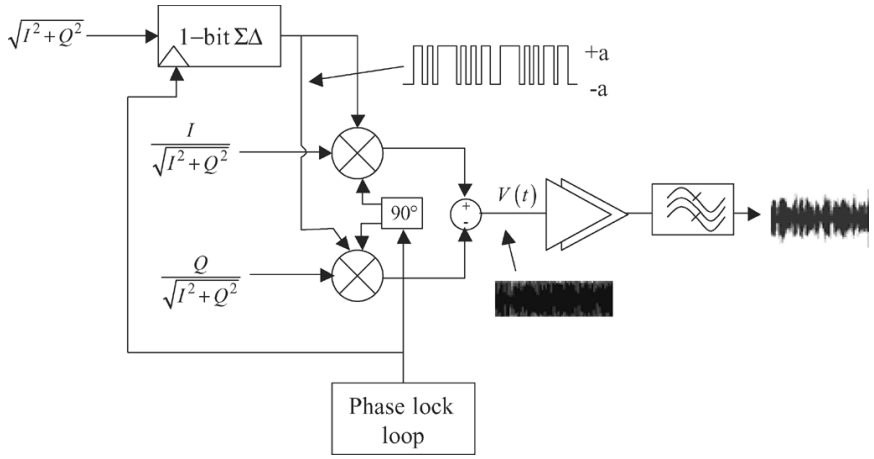


Fig. 3. New transmitter architecture with a 1-bit $\Sigma\Delta$ modulator.

obtained after inverse fast Fourier transform can be written as follows:

$$s(t) = \sum_k \sum_n a_{k,n} \cos(\omega_n t + \varphi_{k,n}). \quad (1)$$

The OFDM signal presents large magnitude variations, as demonstrated with 64 subcarriers in Fig. 4, since the difference between the maximum and the mean power is up to 18 dB, while this difference is only of 6 dB for the classical 16 QAM realized with a raised cosine filter with a 0.5 roll off.

As a consequence, this modulation is highly sensitive to power amplifier AM/AM and AM/PM distortions. Bandwidths of the envelope and phase signals are also large, the former above 60 MHz and the latter above 80 MHz [6]. Such characteristics would tighten the transmitter's specification and so point up its limitations in comparison with the 16 QAM.

III. Transmitter Architecture With Classical PWM Generation

A. PWM Envelope Coding

The first method used for the PWM generation is based on the comparison between the signal and a sawtooth or triangle shape as presented in Fig. 5.

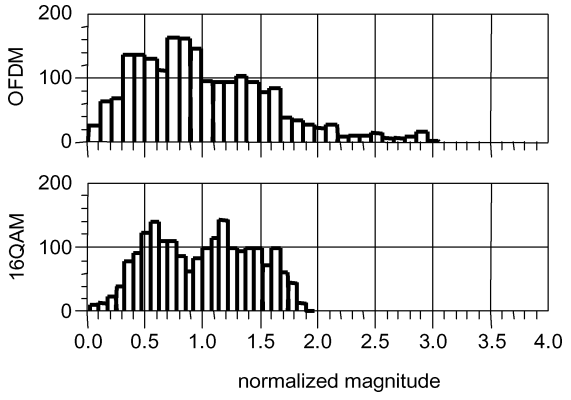


Fig. 4. Envelope magnitude variation for a 64 sub-carrier OFDM in comparison with a 16 QAM.

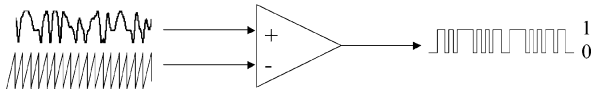


Fig. 5. PWM using a sawtooth signal and an operational amplifier.

The output signal is then a series of pulses with "0" or "1" values. The result of its multiplication with the phase signal, see

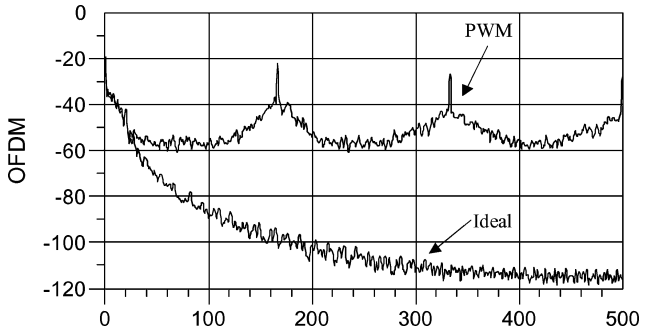


Fig. 6. Power spectral density of the PWM coded OFDM envelope using a 180-MHz saw_tooth reference signal by comparison with the ideal OFDM envelope signal PSD.

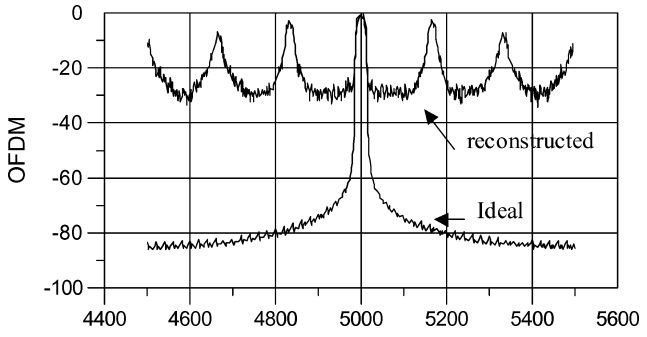


Fig. 7. Ideal and reconstructed signal spectrums for OFDM modulation.

Fig. 2, is then amplified for instance with a Class-C or Class-E switched amplifier.

An important drawback of this method is the high level of intermodulation terms. When applied to an envelope signal, this method generates harmonics of the reference signal and mixing products as presented in Fig. 6. For both modulation schemes, the optimum shape and frequency of the reference signal have to be found to reduce spurious levels. These optima are selected as an heuristic tradeoff between the noise level, the locations of peaks and the transmitter EVM. They are consistent with theoretical results of [7]. For the 16 QAM, the optimum shape is a 200-MHz sawtooth, whereas for the OFDM modulation, although the optimum shape is the same, the frequency was lowered to 180 MHz. Increasing the reference frequency enlarges intermodulation terms and degrades transmitter EVM (8% with a 1-GHz sawtooth).

B. Transmitter Performances

Fig. 7 presents the signal spectrum at the input of the power amplifier obtained from the multiplication of the PWM envelope signal with the phase signal and the local oscillator. We find the spectrum of Fig. 6, shifted in frequency and reconstructed with the phase signal so that the central part is an real OFDM signal.

This method raises two difficulties: the output filter bandwidth and the EVM value.

As far as the first difficulty is concerned, radio-communication standards always require spectrum template and low level of spurious emissions. With the presence of mixing products and high reference harmonics, spurious emissions out of the channel bandwidth are at least -30 dBc in 1-MHz bandwidth.

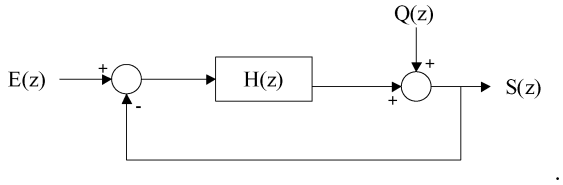


Fig. 8. Linearized model of a $\Sigma\Delta$ modulator.

If the standard specifies that the level has to be lower, the output filter would have to have a low relative bandpass: in our example this corresponds to 0.4% for a 5-GHz carrier. The second difficulty concerns the value of the ideal EVM value: 1.5% for the 16 QAM but 3.2% for the OFDM modulation. For the latter, when a Class-E power amplifier is added into the simulation, the EVM reaches 4% but with no degradation of the output spectrum.

With additional component imperfections, such as I/Q modulator mismatch and synthesizer phase noise both impacting EVM of about 5%, the global EVM value is too high for radio-communication systems [6]. This implies that the limits of this solution have been reached.

IV. Transmitter Architecture Using a 1-Bit $\Sigma\Delta$ Modulator

In this section, we show that the use of a "$\pm a$ $\Sigma\Delta$" ends with a constant envelope signal at the input of the PA. This key point will result in a highly linear and efficient transmitter.

A. Constant Envelope Property

An alternative to classical PWM coding is the use of a 1-bit $\Sigma\Delta$ modulator, see Fig. 3, in which the quantified noise is moved out of band. The aim of its design is then to reject the noise out of the envelope bandwidth in such a way that it releases the output filter bandwidth. After the multiplication between the 1-bit $\Sigma\Delta$ output, the phase signals and the local oscillator, we obtain:

$$\begin{aligned} V(t) &= \pm\, a.\,(A\cos(\omega t)\cos\varphi - A\sin(\omega t)\sin\varphi) \\ &= \pm\, a.A\cos(\omega t+\varphi) \\ &= Aa\cos(\omega t+\varphi\pm\pi) \end{aligned} \tag{2}$$

which is the expression of a constant envelope signal.

In comparison with the former PWM solution or with the "0/1 $\Sigma\Delta$" solution of [4], the power amplification is highly simplified so that any type of amplifier can fit this transmitter since its input signal has a constant envelope (the 0/1 solution may require a specifically designed linear amplifier optimized for establishment time).

B. $\Sigma\Delta$ Architecture and Envelope Coding

The linearized model of a 1-bit $\Sigma\Delta$ is given in Fig. 8.

The output signal can then simply be written as follows:

$$S(z) = \frac{H(z)}{H(z)+1}E(z) + \frac{1}{H(z)+1}Q(z) \tag{3}$$

with $Q(z)$ the quantization noise.

The first term of this equation corresponds to the signal transfer function (STF) and the second part corresponds to the

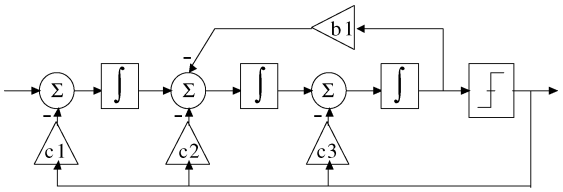


Fig. 9. AΣΔ feedback structure with optional local feedback.

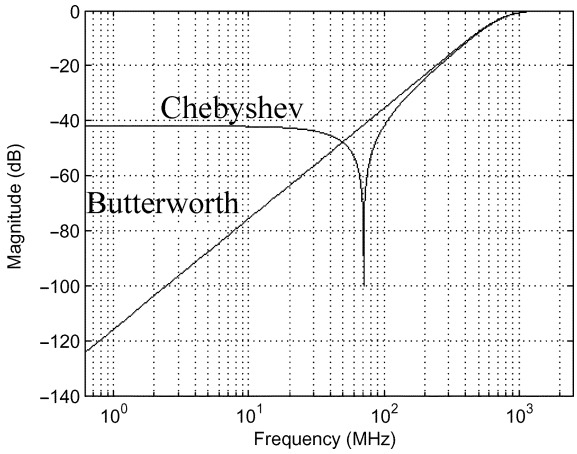


Fig. 10. Theoretical NTF of 1-bit second-order ΣΔ with Butterworth and inverse Chebyshev structures.

TABLE I
ΣΔ SUMMARIZED PERFORMANCES

	2nd Butt	2nd Cheb	3rd Butt	3rd Cheb
NTF gain	1.93	1.95	1.61	1.69
Input level (dB)	-11.8	-11.4	-14.5	-15.5
SNR (dB)	44.6	46.8	46.5	54.0
Noise (dB) at 100MHz	-70	-80	-70	-90

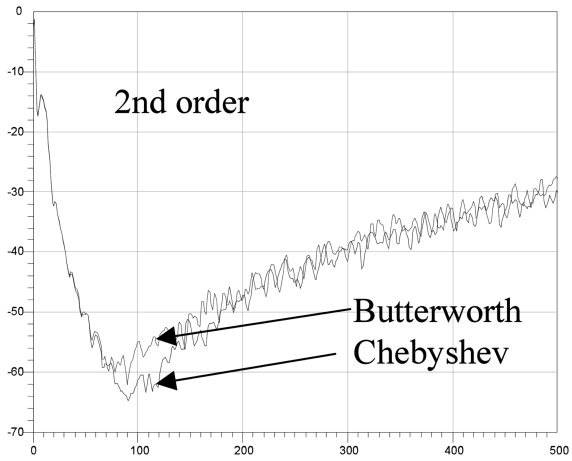


Fig. 11. 1-bit second-order ΣΔ modulator output spectra with an OFDM envelope signal as an input for Butterworth and Chebyshev structure.

noise transfer function (NTF), both dependent on $H(z)$. In the architecture, the ΣΔ is devised so that the former is a low pass filter and the latter a high pass one. This is achieved with a structure such as the one in Fig. 9, where, depending on the presence of the local feedback b1, a third-order Butterworth or inverse Chebyshev NTF can be obtained.

In our application, we aimed to limit the noise floor at the transmitter output below −40 dBc, in 1-MHz measurement bandwidth, for an output bandwidth of 200 MHz. In order to reach this objective, we calculated several ΣΔ with a minimum bandwidth of 100 MHz. Such a bandwidth can only be realized with a high sampling frequency and to achieve this, we took advantage of the local oscillator to drive the ΣΔ.

As the sampling frequency and the bandwidth are imposed, the oversampling ratio (OSR) is

$$\mathrm{OSR} = \frac{F_s}{2 * Bw} = \frac{5.10^9}{200.10^6} = 25 \qquad (4)$$

With the OSR fixed by the architecture, the only remaining parameters which can be varied are the order and structure of the ΣΔ [8]. The determination and choice of this order and structure must take into account complexity of the design, stability, and the signal to noise ratio. In order to examine these points, four 1-bit ΣΔ were studied: two second-order ones, whose theoretical NTF are presented in Fig. 10, and two third-order ones, all with Butterworth or inverse Chebyshev structure. The major difficulty of their design concerned stability. The structure coefficients were established by adjusting them to the theoretical high pass transfer function needed for the application. Nevertheless, according to the value of the linearizedquantizer gain, which depends on the input signal level, instability may occur. This was evaluated using a root locus technique [9] that enables the specification of the maximum stable input signal level. A second criterion to avoid instability is to keep the ΣΔ NTF gain below 2 [10].

Table I summarizes the four ΣΔ performances, in the case of a sinusoidal input, with the following parameters: NTF gain, the input signal level relative to a feedback signal equal to 1, the SNR in 100-MHz bandwidth and the noise level at 100 MHz (relative to a feedback signal equal to 1).

For all ΣΔ, the noise level at $F_s/2$ remain under −20 dB relative to a feedback signal equal to 1.

As expected, the SNR increases between the Butterworth and inverse Chebyshev structures. This gain is about 2 dB for a second-order ΣΔ and more than 7 dB for the third-order one. However, when the two Butterworth structures are compared, the SNR amelioration is insufficient in view of the increase in the complexity of the structure.

The previous analysis was done with a pure sinusoidal input. Therefore, results and conclusions may be quite different with ODFM or 16-QAM envelope signal. The envelope signal is applied in such a way that its mean level corresponds to the equivalent sinusoidal signal indicated in Table I so that the relative maximum input level never exceeds the maximum stable signal. Fig. 11 presents 1-bit ΣΔ modulator output spectra with the ODFM modulation for the two second-order configurations. The SNR improvement between the two structures, whatever the order, is about 5 dB. Likewise, the same result is found when comparing the second and third order.

TABLE II
RMS EVM PERFORMANCES

	2^{nd} Butt	2^{nd} Cheb	3^{rd} Butt	3^{rd} Cheb
16 QAM	0.46%	0.52%	1.78%	1.36%
OFDM	0.35%	0.40%	1.26%	0.78%

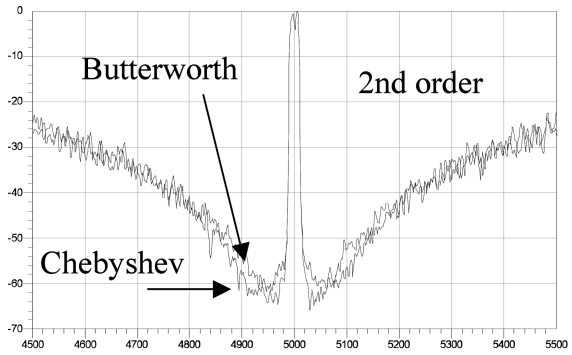


Fig. 12. Output spectra with an OFDM modulated signal for second-order Butterworth and Chebyshev $\Sigma\Delta$.

C. Transmitter Performances

The real impact of the $\Sigma\Delta$ choice has to be evaluated on the global transmitter architecture after the bandpass filter. In our simulations, to avoid distortions linked to variations of propagation time, we chose a raised cosine filter with 200-MHz bandwidth.

The output signal is formed by the multiplication of envelope and phase signals, which corresponds to a convolution in the spectral domain. Although higher order $\Sigma\Delta$ have better noise performances inside the signal bandwidth, the out of band noise shape presents a steeper slope. As a result, the convolution between the phase signal and the out of band noise produces a deterioration in the SNR which more than offsets the improvement linked to $\Sigma\Delta$ order increase, as demonstrated in Table II with rms EVM values.

In comparison with classical PWM architecture, the EVM results are better: 0.35% with OFDM modulation and second-order Butterworth $\Sigma\Delta$ instead of 3.2%. Concerning the "0/1" $\Sigma\Delta$, simulations for OFDM modulation showed a high EVM ($> 11\%$) due to bad envelope restoration.

The $\pm a$ $\Sigma\Delta$ architecture using an OFDM modulation was also simulated with a Class-B power amplifier and gives an EVM value of 0.38% with just slight modifications of the output spectrum.

It is also important to note that this architecture is less sensitive to the modulation type than the previous one. Moreover, the output spectra of OFDM and 16 QAM are very similar and both respect the -40 dBc constraint in a 200-MHz bandwidth. Concerning the spectrum to be filtered, see Fig. 12, the absence of harmonics makes it easier to filter and generates less disturbance on the radio part. The output filter would have to keep its filtering property on a large bandwidth. As the quantization noise is at the input of the power stage, the SNR remains constant whatever the output power.

The efficiency of this architecture relies mainly upon the $\Sigma\Delta$ consumption and the output power. At high frequency, the $\Sigma\Delta$ will consume few hundred milliwatts (a 4-GHz $\Sigma\Delta$ in [11] consumes 350 mW). In the Hiperlan2 case, PA are usually used at 6 dB back-off (Anadigics, RFS P5022), which decreases their efficiency from 56% to 14%. For a 1-W transmitter, the architecture presented could reach 45% efficiency with the $\Sigma\Delta$ in [11] and the same PA working at its 1-dB compression point. Clearly, using a more efficient PA such as Class E, the efficiency will be even greater.

V. CONCLUSION

This paper presents a new versatile digital transmitter architecture that can be adopted for any non constant modulation, on a wide range of frequencies, with high linearity and efficiency. This architecture is an improvement of the classical EER architecture, and consists in PWM coding the envelope of a modulated signal. Two methods for the PWM generation are analyzed: the simplest method transforms the modulated signal into a switched one. This generates harmonics and intermodulation terms, so this architecture may present insufficient performances for some modulation schemes. The second solution uses a "$\pm a\,\Sigma\Delta$" modulator and enables the transformation of any modulated signal into a constant envelope one. This requires a high frequency 1-bit $\Sigma\Delta$ modulator that can be devised either in digital CMOS or BiCMOS technologies, according to the sampling frequency.

Even with severe modulations such as OFDM, this solution gives the desired level of performance in terms of EVM and output spectra. The other advantage compared to classical EER concerns the dc–dc converter which would be easier to implement.

REFERENCES

[1] L. Kahn, "Single sideband transmission by envelope elimination and restoration," in *Proc. IRE* , Jul. 1952, pp. 803–806.
[2] S. Krishnan and D. Scott *et al.*, "An 8-GHz continuous-time $\Sigma - \Delta$ analog-digital converter in an InP-based HBT technology," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 12, Dec. 2003.
[3] J. C. Jensen and L. E. Larson, "A 16-GHz Yltra-high-speed Si-SiGe HBT comparator," *IEEE J. Solid-State Circuits*, vol. 38, no. 9, pp. [AU: PAGE NUMBERS?]–, Sep. 2003.
[4] Y. Wang, "An improved Kahn transmitter architecture based on delta–sigma modulation," in *Proc. IEEE Microw. Symp.*, Philadelphia, PA, Jun. 2003.
[5] C. Berland, J. F. Bercher, I. Hibon, M. Villegas, D. Belot, D. Pache, and V. Le Goascoz, "Dispositif de modulation et émetteur comprenant un tel dispositif," French Patent FR0 405 636, May 25, 2004.
[6] G. Baudoin, C. Berland, M. Villegas, and A. Diet, "Influence of time and processing mismatches between phase and envelope signals in linearization systems using envelope elimination and restoration, application to Hiperlan2," in *Proc. IEEE Microw. Symp.*, Philadelphia, PA, Jun. 8–13, 2003.
[7] E. R. Kretzmer, "Distortion in pulse-duration modulation," *Proc. IRE*, pp. 1230–1235, Nov. 1947.
[8] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters. Theory, Design, and Simulation*. New York: IEEE Press, 1997, ch. 3.
[9] A. Gothenberg, B. Li, and H. Tenhunen, "A method for stability and performance analysis of low oversampling ratio higher order sigma–delta noise shaper architectures," in *Proc. IEEE 42nd Midwest Symp. Circuits Syst.*, vol. 1, 1999.
[10] K. C. H. Chao, S. Nadeem, W. L. Lee, and C. G. Sodini, "A higher order topology for interpolative modulators for oversampling A/D converters," *IEEE Trans. Circuits Syst.*, vol. 37, no. 3, pp. 309–318, Mar. 1990.
[11] W. Gao, J. A. Cherry, and W. N. Snelgrove, "A 4-GHz fourth-order SiGe HBT band pass $\Delta\Sigma$ modulator," in *Dig. Tech. Papers Symp. VLSI Circuits*, Jun. 11–13, 1998, pp. 174–175.

8.10 C. Joubert, J.-F. Bercher, and G. Baudoin, « Contribution to the study of a Phase-Domain ADPLL », version longue de l'article , "Contributions to the analysis and design of an ADPLL," in *13th IEEE International Conference on Electronics, Circuits and Systems, 2006. (ICECS) 2006.*, 2006, pp. 322–325.

Contribution to the study of a Phase-Domain ADPLL

Cyril Joubert(1), Jean François Bercher(1), Geneviève Baudoin(1)

(1) ESIEE / ESYCOM, 2 Boulevard Blaise Pascal, 93160 Noisy-le-Grand, France.

***Abstract* — In this paper, we present contributions to the analysis and implementation of a Phase-Domain All-Digital Phase-Locked Loop (ADPLL) for RF applications. First, we propose a Time Behavioral Model of the ADPLL including the modelization of the necessary dithering effects. Second, we present a baseband analysis that allows to compute the power spectral density from the instantaneous frequency obtained as the output of the behavioral model. Third, we propose and interpret adaptive algorithms that can be integrated in the ADPLL in order to lower its complexity. These different aspects are illustrated through simulations.**

I. INTRODUCTION

Staszeswski *et al.* recently presented a new All Digital Phase-Locked Loop based RF frequency synthesizer [1].

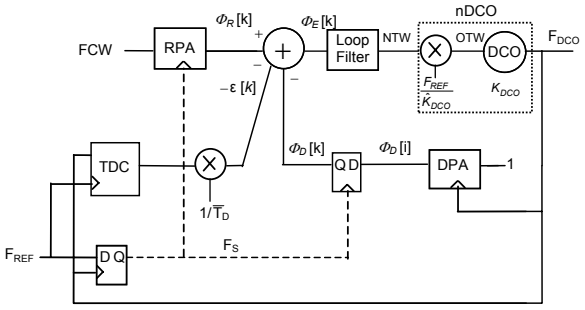


Fig 1. ADPLL based RF frequency synthesizer [1]

A digitally controlled oscillator (DCO) allows for this PLL to be implemented in a fully digital manner [2].

Phase accumulators are used to count cycle periods of reference and feedback oscillators. A synchronous clock, F_S, undersamples the output of the DCO phase accumulator (DPA), so that comparison of the two phases can be performed using the same clock. The retimed clock, F_S, is achieved by oversampling the reference clock, F_{REF}, by the oscillator clock, F_{DCO}. Note that in Fig 1, index *i* and *k* do not refer to the same clock. Also, the DCO is normalized using a compensation gain, in such a way that its input the normalized Tunning word (NTW), becomes independent of the gain K_{DCO} of the oscillator.

ADPLL precision depends on phase detector performance and DCO resolution. DCO resolution can be improved using dithering as will be presented later. Concerning the first aspect, higher ADPLL precision can be obtained using fractionnal phase error correction. The Time to Digital Converter (TDC) is used to convert the delay between the RF and reference clocks directly into a digital quantity [3], with a time resolution, noted ΔT_{RES}, that can be equal to the elementary propagation delay through an inverter gate. This delay is converted to a normalized phase using the normalization $1/\overline{T}_D$, where $\overline{T}_D$ is an average value of the DCO period..

The Frequency Command Word (FCW) is given as input to the reference phase accumulator (RPA), and enables to tune the output frequency of the DCO.

$$F_{DCO} = FCW \times F_{REF} \qquad (1)$$

Designing a PLL requires a simulator in order to study the effect of varying parameters and optimize the PLL.

Analysis and simulation of the ADPLL in Fig. 1, with a direct method requires a very high rate clock. Indeed, such a clock must have a rate that is greater than the highest frequency in the system. This requirement leads to an incredible simulation time and fantastic amounts of data.

The objective of this paper is to show that it is possible to simulate such a PLL at much more reasonnable rate, without sacrifying accuracy and flexibility. We propose here a directly implementable model that also allow easy access to all variables of interest.

Our behavioral model is presented in sections II to IV. Part of the material in section II and III was already introduced in the RWS 2006 conference [7]. In section III, the model of the TDC is added and in section IV, the whole behavioral model is significantly extended in order to include the Sigma-Delta modulator. The performance of the PLL is directly linked to the quality of the Power Spectral Density at its output. In section V, we recall, as in [7], that the PSD can be computed easily as a by-product of the low-pass simulation. Implementation of the PLL can be simplified using adaptive algorithms instead of harware complexity. In section VI, we argue and work out two adaptive algorithms. Finally, all these results are illustrated in section VII.

II. PRINCIPLE OF THE BEHAVIORAL MODEL

The limiting components are the DCO phase accumulator (at F_{DCO} rate), and the TDC (with accuracy ΔT_{RES}), because they would need an extremely fast sampling frequency for a correct representation. The key of the new simulation model is the fact that outputs of limiting components can be expressed directly at rate F_{REF} (this can be understood as a kind of carrier frequency suppression).

In order to understand the principle of the behavioral model, we focus on simple D flip-flop. Indeed, this central element links asynchronous clocks F_{DCO} and F_{REF} because it resynchronizes them.

Let us consider the waveform shown in Fig 2. In this diagram two important parameters appear: the delay τ_k and the integer value *N(k)*. τ_k is defined as the difference between the k^{th} reference rising edge and following oscillator rising edge, and *N(k)* is the real-value count of the DCO clock periods $T_{D,k}$ for each cycle of the reference clock.

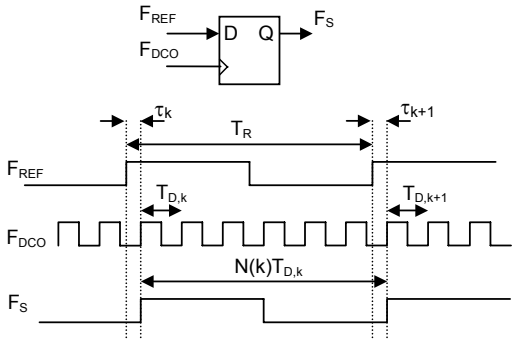


Fig 2. Inputs/output of the D flip-flop with a representative set of waveforms

The exact relation between reference and oscillator frequencies can be deduced from the waveform in Fig 2. We obtain the following relashionship between τ_{k+1} and τ_k

$$\tau_{k+1} = \tau_k + N(k)T_{D,k} - T_R \qquad (2)$$

Equation (2) is valid if $T_{D,k}$ remains constant during a whole cycle of F_{REF}.

An important point is that the time delay is bounded according to

$$0 \le \tau_{k+1} \le T_{D,k} \qquad \forall k \qquad (3)$$

Another essential remark is that the phase error in the PLL is directly proportionnal to this delay.

Let us note by $\lfloor x \rfloor$ the integer part of x and define by $N_i(k)$ the integer part of the ratio between the two periods acoording to:

$$N_i(k) = \lfloor T_R / T_{D,k} \rfloor \qquad (4)$$

In (2) we have either $N(k) = N_i(k)$ or $N(k) = N_i(k) + 1$. Hence the behavior of the D flip-flop is equivalent to a Dual Modulus Divider (DMD) controlled by the phase error (via the delay τ) as shown in Fig 3.

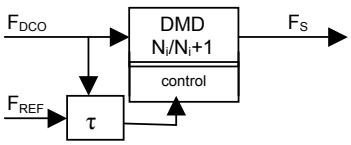


Fig 3: Behavioral model of the D flip-flop by a DMD controlled by the phase error

In the case where *N(k)* takes only two values, computation of *N(k)* proceeds as follows:

First suppose that $N(k)=N_i(k)+1$, and compute τ_{k+1} using (2). Then, we have to check that (3) is satisfied. If it is, we keep $N(k)=N_i(k)+1$, otherwise $N(k) = N_i(k)$.

This analysis results in the following equations:

$$\tau(k+1) = \begin{cases} (N_i+1)T_{D,k} - T_R + \tau(k) & if \quad \tau(k+1) - T_{D,k} < 0 \\ N_i T_{D,k} - T_R + \tau(k) & otherwise \end{cases} \qquad (5)$$

This can be further simplified into

$$\tau(k+1) = \tau(k) + \left(N_i(k) + 0{,}5\right)T_{D,k} - T_R - \operatorname{sgn}\left(N_i(k)T_{D,k} - T_R + \tau(k)\right)T_{D,k}/2 \qquad (6)$$

using sgn(x) the sign function: sgn(x)= –1 if x<0, and sgn(x)= 1 otherwise.

Similarly, equation (2) and condition (3) leads to

$$N(k) = N_i(k) + 0.5 - 0.5\operatorname{sgn}\left(\tau(k) + N_i(k)T_{D,k} - T_R\right). \qquad (7)$$

III. Closed Loop Model

For the closed loop model we need to compute the output of phase accumulators and fractionnal error correction ε. Expression of reference phase accumulator is given by

$$\phi_R[k+1] = \left(\phi_R[k] + FCW\right) \bmod [2^R] \qquad (8)$$

Similarly, the undersampled output of the DCO phase accumulator, $\phi_D[k]$ of the finite width *D* can be written

$$\phi_D[k+1] = \left(\phi_D[k] + N[k]\right) \bmod [2^D] \qquad (9)$$

However this expression cannot be implemented without knowledge of *N(k)*. Thanks to our previous analysis, we are here able to compute *N(k)* using (7) and, therefore implement (9).

The fractionnal phase error ε can be simply modelized as the quantified version of our previous τ_{k+1}, normalized to $\bar{T}_D$, an averaged value of T_D.

$$\varepsilon(k) = \left\lfloor \frac{\Delta T_{RES}}{\bar{T}_D} \left\lfloor \frac{\tau(k+1)}{\Delta T_{RES}} \right\rfloor 2^{R-D} \right\rfloor \qquad (10)$$

A more precise model can be derived from the analysis of the TDC in terms of quantified delay between rising (ΔT_R) and falling (ΔT_F) edges preceding the rising edge of F_{REF} [3] (cf. Fig 4).

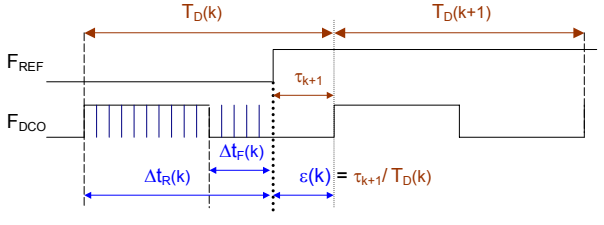


Fig 4. Inputs / outputs of the TDC

The time-to-digital conversion [3] is realized by passing the DCO signal through a chain of inverters gates of typical delay ΔT_{RES}. Then, each delayed outputs are sampled by the same reference clock. The Edge Detector detect rising and falling edge transition and realize a pseudo-thermometer encoding to deliver the numerical data ΔT_R and ΔT_F.

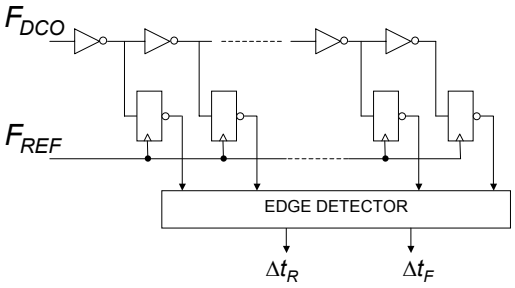


Fig 5. Time-to-Digital Converter (TDC) architecture

The outputs of the TDC are modelized by

$$\Delta t_R(k) = \left\lfloor \frac{T_D(k) - \tau(k+1)}{\Delta t_{RES}} \right\rfloor$$
$$\Delta t_F(k) = \begin{cases} \left\lfloor \frac{T_D(k)/2 - \tau(k+1)}{\Delta t_{RES}} \right\rfloor & if \quad \tau(k+1) < T_D(k)/2 \\ \left\lfloor \frac{3.T_D(k)/2 - \tau(k+1)}{\Delta t_{RES}} \right\rfloor & otherwise \end{cases} \quad (11)$$

And the fractionnal phase error is finally estimated by

$$\varepsilon(k) = 1 - \frac{\Delta T_R(k)}{\overline{T}_D} \quad (12)$$

with $\overline{T}_D$ an average value of $T_D(k)$ used for normalized the TDC output (cf. Fig 1).

$$\overline{T}_D = \frac{1}{N_{AVG}} \sum_{k=1}^{N_{AVG}} 2\left|\Delta T_R(k) - \Delta T_F(k)\right| \quad (13)$$

An LMS algorithm is proposed in section VI-B in order to compute efficiently the inverse of $\overline{T}_D$ involved in (11).

The phase error at the input of the loop filter in Fig 1, is computed in digital signed format according to

$$\phi_E(k) = \phi_R(k) + \varepsilon(k) - \phi_D(k).2^{R-D} - 2^R\left(\left\lfloor \frac{\phi_R(k) + \varepsilon(k)}{2^{R-1}} \right\rfloor - \left\lfloor \phi_D(k).2^{-D+1} \right\rfloor\right) \quad (14)$$

The DCO is modelized by equations given in [2]. For a small deviation Δf, we can use simple linearized model

$$f_{DCO}(k) = f_0 + \Delta f(k) = f_0 + OTW(k)K_{DCO} \quad (15)$$

where f_0 is the central frequency, OTW is the oscillator tuning word at the input of the DCO, and K_{DCO} the gain of the DCO.

Note that the actual output of our model is directly the instantaneous frequency, delivered at rate F_{REF}, and not a time signal with that instantaneous frequency. However, we may compute the PSD of such virtual signal as described in section V.

IV. Account for the Sigma Delta Modulator

A. ADPLL Locking sequence

Convergence of the ADPLL is achieved using three different modes [2]: first, a calibration (CAL) mode initiates the TDC and the central frequency of the PLL, independantly of frequency command word. Second, an acquisition (ACQ) mode acquires the channel selected by FCW. Third, a tracking (TRK) mode achieves the required performance (use of a Sigma Delta modulator refines this last mode).

In Fig 6, for each mode, the normalized tuning word (NTW) is converted and resynchronized. The oscillators tuning words OTW are connected to different capacitor banks in the DCO.

One can note that the behavioral model developped above is adequate for analysing the Calibration and Acquisition modes, but is no more valid for the Tracking mode, because the rate, due to the ΣΔ modulator, at the input of the DCO is much higher than F_{REF}.

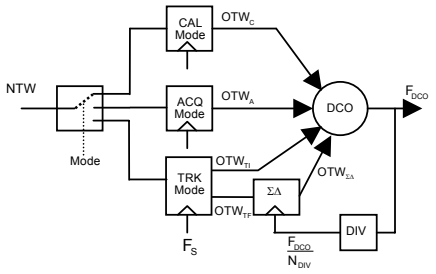


Fig 6. Normalized DCO scheme

Indeed, in (2), the hypothesis that $T_{D,k}$ remains constant during a whole cycle of F_{REF} is no more true. Thus the outputs *N(k)* and τ_k of our behavioral model must be rederived so as to take this into account.

B. Input/Output Time Behavioral Model of the ΣΔ

The first step is to express the input/output behavior of the ΣΔ. A simple m^{th} order ΣΔ modulator model can be obtained by the following equation [4],

$$y(z) = x(z) - \left(1 - z^{-1}\right)^m r(z) \quad (16)$$

where x is the input, y the output and r is a white quantization noise uniformly distributed between 0 and 1. A more precise model can be derived from the purely digital MASH structure [5] implemented in this ADPLL.

C. Closed loop model with ΣΔ

Our objective is to rederive the delay τ_k and the integer value *N(k)* defined in the section II.

The preceding model used only the rate F_{REF}, defined by the index k, whereas ΣΔ modulator now uses a new rate at $Fdco/N_{DIV}$, with a temporal index denoted by i.

At the end of each F_{REF} cycle, we need to compute two secondary parameters: $\tau_{\Sigma\Delta,k+1}$, which is represented in Fig 7, and $N_{\Sigma\Delta}(k)$ and integer which is calculated as follows.

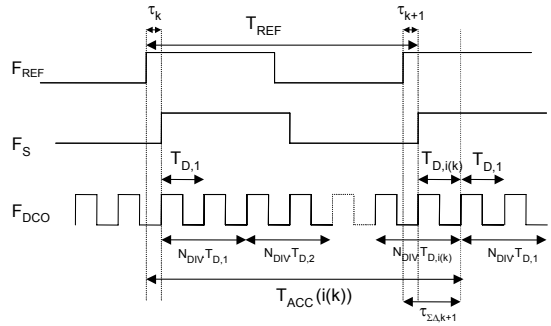


Fig 7 : Representative set of waveforms for ΣΔ behavioral model with $N_{DIV} = 2$.

For each Fref cycle k, we define at each instant i the accumulated time T_{ACC} by

$$T_{ACC}\left(i(k)\right) = \tau_k + N_{DIV}\sum_{m=1}^{i(k)} T_{D,m} \qquad (17)$$

where $T_{D,i(k)}$ is the period of the DCO that may change at the ΣΔ rate. $T_{D,i(k)}$ is calculated thanks to (16) and (17).

When $T_{ACC}(i(k)) > T_{REF}$ the secondary parameters are calculated by

$$\begin{aligned} \tau_{\Sigma\Delta,k+1} &= \mathrm{T_{REF}} - T_{ACC}\left(i(k)\right) \\ N_{\Sigma\Delta,k+1} &= \left\lfloor \frac{\tau_{\Sigma\Delta,k+1}}{T_{D,i(k)}} \right\rfloor \end{aligned} \qquad (18)$$

and we deduce the principal parameters of this model by

$$\begin{aligned} \tau_{k+1} &= \tau_{\Sigma\Delta,k+1} - N_{\Sigma\Delta,k+1} T_{D,i(k)} \\ N(k) &= N_{DIV}.i(k) + N_{\Sigma\Delta,k} - N_{\Sigma\Delta,k+1} \end{aligned} \qquad (19)$$

Thus we can keep the previous model for the tracking mode, simply by refining it by the incorporation of the two new equations (18) and (19). Actually, only (15) is evaluate at the ΣΔ rate, while (16) and (17) are only evaluated at rate F_{REF}. Therefore, the complexity does not increase much, and remains several orders less than a conventional "high rate" simulator.

V. Spectral density computation

The objective of this section is to show how to compute the PSD from the instantaneous frequency given by the output of the behavioral model.

Let us consider f_i(t) the instantaneous frequency of an oscillator

$$f_i(t) = f_0 + \Delta f_i(t) = f_0 + \Delta f_{pp}\, g(t) \qquad (20)$$

where f_0 is the mean frequency, Δf_{pp} is the peak-to-peak deviation from f_0 and $g(t)$ is a normalized frequency modulation pattern ($-1 < g(t) < 1$) with zero mean.

We define the instantaneous phase by

$$\theta_i(t) = \int_0^t 2\pi f_i(\tau)d\tau = 2\pi f_0 t + 2\pi\Delta f_{pp}\int_0^t g(\tau)d\tau. \qquad (21)$$

The output signal $s(t)$ of the oscillator is given by

$$s(t) = \cos\left(\theta_i(t)\right) = \cos\left(2\pi f_0 t + 2\pi\Delta f_{pp}\int_0^t g(\tau)d\tau\right). \qquad (22)$$

With the assumption that $x = 2\pi\Delta f_{pp}\int_0^t g(\tau)d\tau$ is small enough so that $\sin(x) \approx x$, we obtain the first order approximation

$$\hat{s}(t) = \cos(2\pi f_0 t) - 2\pi\Delta f_{pp}\sin(2\pi f_0 t)\int_0^t g(\tau)d\tau \qquad (23)$$

after developing the cosine in (22).

The Fourier Transform of $\hat{s}(t)$ is

$$\hat{S}(f) = \frac{1}{2}\left(\delta_{f0} + \delta_{-f0}\right) + \frac{\Delta f_{pp}}{2}\left(\frac{G(f-f_0)}{f-f_0} - \frac{G(f+f_0)}{f+f_0}\right) \qquad (24)$$

where $G(f) = TF\left[g(t)\right]$.

For $f = \pm f_0$, the weight can also be computed by $1/2 - \pi\Delta f_{pp}\frac{1}{t_{\max}}\int_0^{t_{\max}}\int_0^t g(\tau)d\tau\, dt$.

Thereby, with (24), we can directly compute the PSD transform of the instantaneous frequency using solely the lowpass signal $g(t)$. The analysis above is continuous, but a similar analysis can be done in the discrete case. For implementation of a simulator based on our behavioral model, care must be taken on spectral aliasing, and a zeroth-order interpolation have to be used in order to increase the sampling period.

VI. LMS algorithms for adaptative estimate

In the all-digital PLL, one can simplify the architecture using adaptive algorithms instead of higher hardware complexity.

A. DCO compensation gain estimation

In a forthcoming paper [6], Stazsewski *et al.* propose a LMS adaptation algorithm in this direction. In order to estimate the normalization gain $\hat{K}_{DCO}^{-1}$, they present the simple adptation rule:

$$\hat{K}_{DCO}^{-1}[n] = \hat{K}_{DCO}^{-1}[n-1] + \mu\nabla \qquad (25)$$

with a sign algorithm, $\nabla = \phi_E\, sign(FCW)$. We show here that such an algorithm can be interpreted as the minimization of a simple criterion.

Indeed, when the PLL is settled, if we apply a ΔFCW step on the frequency command word (FCW), we obtain a deviation of the phase error $\Delta\phi_E = (1-r)\Delta\phi_m$, with $r = K_{DCO}/\hat{K}_{DCO}$, and $\Delta\phi_m$ the phase deviation related to ΔFCW.

Note that this reasoning is correct only for a type I PLL or when ΔFCW is applied as a 2 points modulation scheme.

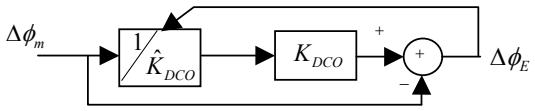


Fig 8. Adaptative algorithm for identification of gain compensation

Then, using the minimization criterion $J(r) = \mathrm{E}\left[\left|\Delta\phi_E\right|^2\right]$ as shematized in Fig 8. The classical adaptation equation of a gradient algorithm $r[n] = r[n-1] - \mu.\nabla J$, with $\nabla J = 2\mathrm{E}\left[\Delta\phi_E\ \partial\Delta\phi_E/\partial r\right]$ and $\partial\Delta\phi_E/\partial r = -\Delta\phi_m$ results in the LMS recursion

$$r[n] = r[n-1] + 2\mu.\Delta\phi_E.\Delta\phi_m\,, \qquad (26)$$

using the instantaneaous estimate of ∇J.

With $r[n] = K_{DCO}[n]/\hat{K}_{DCO}[n]$, we also obtain

$$F_{REF}.\hat{K}_{DCO}^{-1}[n] = F_{REF}\left(\hat{K}_{DCO}^{-1}[n-1] + \frac{2\mu}{K_{DCO}}\Delta\phi_E.\Delta\phi_m\right) \quad (27)$$

Last, with $\mu_0 = 2\mu F_{REF}/K_{DCO}$ and using the sign of the error, we obtain

$$F_{REF}.\hat{K}_{DCO}^{-1}[n] = F_{REF}.\hat{K}_{DCO}^{-1}[n-1] + \mu_0.\phi_E sign(\Delta\phi_m) \quad (28)$$

and with $sign(\Delta\phi_m) = sign(\Delta FCW)$ we recover the proposed algorithm (25).

B. Inverse DCO period estimation

The computation of the fractionnal phase error ε involves the multiplication of the output of the TDC by the inverse period of the DCO, $1/\overline{T}_D$ (cf. Fig 1).

A possible approach [3] is to estimate $\overline{T}_D$ by an average of N_{AVG} values of T_D, and then use a digital divider. As in the case of the DCO compensation gain, this can be simplified using an adaptative method, which results in the inversion without divider, and provides adaptivity to the context.

Let $\alpha = 1/\overline{T}_D$. Then clearly, α can be found as a minimizer of $J(\alpha) = \mathrm{E}\left[\left|1-\alpha\overline{T}_D\right|^2\right]$.

Thus, we can adopt the simple LMS algorithm

$$\alpha^{(n+1)} = \alpha^{(n)} - \mu.\nabla J \qquad (29)$$

with $\nabla J = -2\overline{T}_D\left(1-\alpha\overline{T}_D\right)$, the instantaneous estimate of the criterion $J(\alpha)$.

The optimal step can be chosen as $\mu_{opt} = 1/K.\overline{T}_D^{\,2}$ (K between 2 and 4). Using this adptation step and noting that $\alpha \simeq 1/\overline{T}_D$ at convergence, we may rewrite the algorithm as

$$\alpha^{(n+1)} = \alpha^{(n)} + \frac{\alpha^{(n)}}{K}\left(1-\alpha^{(n)}\overline{T}_D\right) = \alpha^{(n)}\left(1+\frac{1}{K}-\frac{\alpha^{(n)}}{K}\overline{T}_D\right) \quad (30)$$

Complexity can be further reduced using a sign algorithm version:

$$\alpha^{(n+1)} = \alpha^{(n)} + \mu_{opt}\,sign\left(1-\alpha^{(n)}\overline{T}_D\right) \qquad (31)$$

Note also that instead of using a single estimation of $\overline{T}_D$, and in order to preserve adaptivity, we can use a sliding window or an exponential mean such as

$$\overline{T}_D^{(N)} = \beta\overline{T}_D^{(N-1)} + (1-\beta)T_{D,N} \qquad (32)$$

with $T_{D,N}$ the Nth value of T_D and β the forgetting factor.

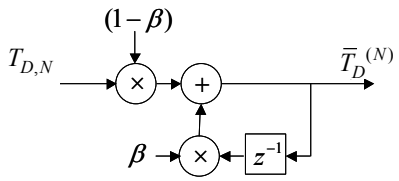


Fig 9. DCO period estimation by exponential mean

VII. SIMULATIONS

The behavioral model was implemented in MATLAB and compared with a conventional model realized in VHDL. We compare both models and present typical results obtained using the behavioral model, give an example of spectral density computation, and illustrate the convergence of the LMS algorithm for T_D inversion.

A. Time behavioral model validation including ΣΔ effects

For both models with same set of parameters, we compare the transient behavior of the DCO output frequency when we change the frequency command word.

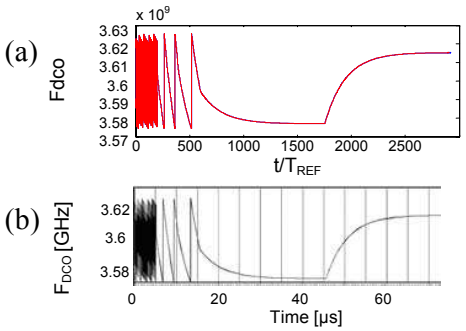


Fig 10. Simulation comparison of instantaneous frequency between new behavioral model (a) and VHDL model (b)

Fig 10 show the transient reponse of the output frequency of the ADPLL, when *FCW* is out of synthesizable range of the DCO (before 16µs) and next, when *FCW* changes between minimum and maximum value of the DCO. We observe the same behavior and settling time.

Fig 11 shows the settling sequence for the ADPLL, using the equations (14)-(15) of the behavioral model.

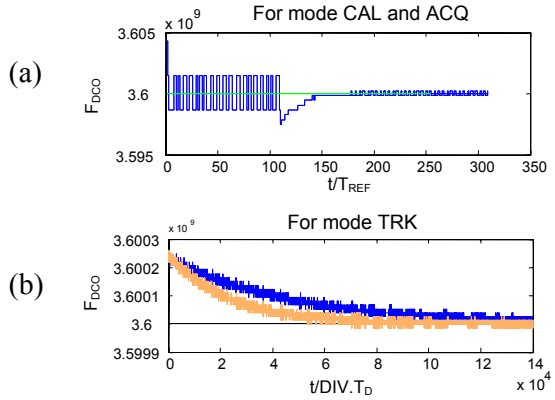


Fig 11. Settling sequence for modes CAL, ACQ (a) and TRK (b) for two different loop filter, and involving the ΣΔ modelization.

B. Spectral density computation

An example of PSD of the instantaneous frequency of the DCO is given in Fig 12.

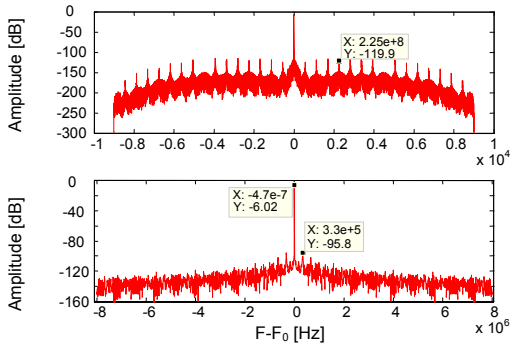


Fig 12. PSD of the locking sequence when PLL is settled

We can note the amplitudes and frequencies of spurs, due to accuracy of TDC, precision of DCO, and ΣΔ effects.

This typical simulation result shows that our simple model can be an useful tool for the development, test and optimisation of an ADPLL.

C. LMS algorithm for T_D inversion

Performances of the LMS algorithm for $\overline{T}_D$ inversion are illustrated in Fig 13. This shows both effectiveness and fast convergence of the algorithm.

Here $\overline{T}_D^{(N)}$ is computed using a sliding window of length 128:

$$\overline{T}_D^{(N)} = \frac{1}{128}\sum_{k=N-127}^{N} 2\left(\Delta t_R(k) - \Delta t_F(k)\right) \qquad (33)$$

Δt_R and Δt_F can be modelized as two uniform random variables, on an interval $\Delta T_{RES} = 13$ ps (the resolution). Then $2\left|\Delta t_R(k) - \Delta t_F(k)\right|$ is distributed according to a triangular distribution, and with a standard deviation of 11.08 ps. By the law of large numbers, $\overline{T}_D^{(N)}$ is then approximatively gaussian with standard deviation of 0.98 ps.

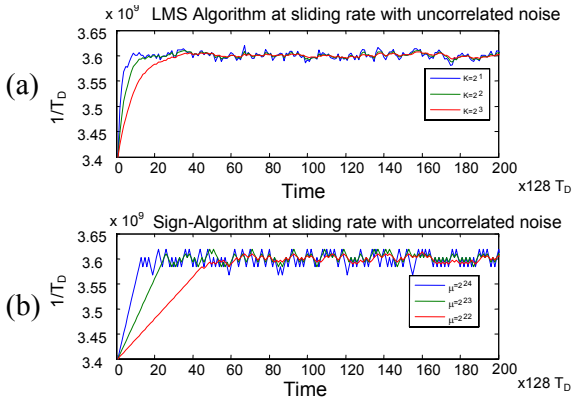


Fig 13. LMS (a) and Sign (b) algorithm for $\overline{T}_D^{(N)}$ inversion using a sliding window

Then when the algorithm is iterated at the "sliding rate", (that is each 128 samples), then the noise is uncorrelated. In this case, it converges easily in less than 20 iterations. Convergence is a bit slower in the correlated case.

VIII. Conclusion

We have presented an effective behavioral model for an all-digital phase locked loop, including dithering effects. This model enables simulating the PLL at low complexity compared to a conventional digital simulator. Furthermore, this simulator enable easy access to all variable of interest. This allowing a study of all parameters of the PLL.

We have also presented and interpreted efficient LMS type algorithms that can lower the hardware complexity.

Further work will include noise sources account in the simulator, analysis of spurious tones, statistical analysis and sensibility of the PLL, and optimization of performances.

References

[1] R. B. Staszeswski and P. T. Balsara, "Phase-–Domain All–Digital Phase–Locked Loop," *IEEE trans. on circuits and System*, vol. 52, no. 3, pp. 159–163, March 2005.

[2] R. B. Staszewski, C–M. Hung, D. Leipold and P. T. Balsara, "A first Multigigahertz Digitally Controlled Oscillator for Wireless Applications," *IEEE Trans. on Microwave Theory and Techniques*, vol. 51, no. 11, pp. 2154–2164, Nov 2003.

[3] R. B. Staszewski, D. Leipold, C–M. Hung and P. T. Balsara, "TDC–Based Frequency Synthesizer for Wireless Applications," *IEEE RFIC Symposium*, pp. 215–218, 2004.

[4] M. H. Perott, M. D. Trott and C. G. Sodini, "A Modeling Approach for Σ–Δ Fractional–N Frequency Synthesizers Allowing Straightforward Noise Analysis," *IEEE J. Solid-State Circuits*, vol. 37, pp 1028-1038, Aug. 2002.

[5] B. Miller and R. J. Conley, "A multiple modulator fractional divider," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 578–583, June 1991.

[6] R. B. Staszewski, J. Wallberg, C-M. Hung, G. Feygin, M. Entezari and D. Leipold, "LMS–based Calibration of an RF Digitally–Controlled Oscillator for Mobile Phones, " accepted for publication in *IEEE Trans on Circuits and Systems II*, April 2005, [Online] http://ieeexplore.ieee.org/iel5/8920/28239/101109TCSII2005858750.pdf?isnumber=28239&arnumber=101109TCSII2005858750.

[7] C. Joubert, J. F. Bercher, G. Baudoin *et al.*, "Time Behavorial Model of Phase domain ADPLL based frequency synthesizer," presented at *RWS 2006*, San Diego, january 2006.

8.11 P. Guillot, P. Philippe, C. Berland, and J. Bercher, "A 2GHz 65nm CMOS digitally-tuned BAW oscillator," in *15th IEEE International Conference on Electronics, Circuits and Systems, 2008. ICECS 2008.*, Sep. 2008, pp. 722–725.

A 2GHz 65nm CMOS digitally-tuned BAW oscillator

P. Guillot, P. Philippe
Innovation Center in RF
NXP Semiconductors
Caen, France
Email: pierre.guillot@nxp.com

C. Berland, J.-F. Bercher
Université Paris-Est
ESYCOM-ESIEE
Noisy-le-Grand, France

***Abstract*— The design of a 2GHz reference frequency oscillator in a 65nm CMOS process using a Bulk Acoustic Wave resonator is presented. The oscillator implements digital frequency control using a switched capacitor bank in parallel to the resonator. The tuning range is up to 4MHz with a minimum step of 1.6kHz. The oscillator core is designed to reach low phase noise (-128dBc/Hz at 100kHz offset) at low power consumption (0.9mW) using a differential topology. It is followed by a low noise divider for output at 500MHz with a phase noise of -140dBc/Hz at 100kHz offset.**

I. Introduction

Wireless communication systems require reference frequency oscillators for synchronization purpose. Conventional frequency references are commonly based on crystal oscillators. They have a high spectral purity but operate at low frequency (26MHz). In addition, a crystal is a rather expensive and is a bulky component that is difficult to integrate in the same package as the RF transceiver. An alternative to crystals can be to use Bulk Acoustic Wave (BAW) resonators. Thanks to their high quality factor and their high resonant frequency in the GHz range, RF oscillators with spectral purity equivalent to crystal oscillators can be achieved [1]. An advantage of BAW resonators is their small size, which, in the future, enables their integration into a single package together with the transceiver and into a single chip radio. In addition to spectral purity, frequency stability is another important concern when considering a reference frequency application. Frequency stability is to a large extent determined by the properties of the resonator. Therefore this question is not directly addressed in this paper where the focus is on the design of the oscillator. However, it is important that the oscillator design provides means for tuning the frequency to the desired value with a high degree of precision and over a frequency range that enables correction of the largest frequency errors.

In this paper, we report a digitally frequency-controlled BAW-based oscillator at 2GHz designed in an advanced 65nm CMOS process. The oscillator has a tuning range of 4MHz with a minimum step of 1.6kHz or 0.8ppm. The targeted application is a reference oscillator. The paper is organized as follows. In section II, a theoretical analysis of the BAW oscillator is presented that illustrates the main design challenges. In section III, the architectures of the oscillator and of the tuning bank are described. Finally, simulation results and comparisons to earlier works are reported in section IV.

II. BAW oscillator analysis

A. Solid Mounted Resonator (SMR)

A BAW resonator consists basically of a piezoelectric layer enclosed between two electrodes. By application of a RF signal on the electrodes, an acoustic wave is excited within the piezoelectric layer. Resonance occurs at a frequency determined by the thickness of the layer. In order to achieve a high quality factor, the wave is confined within the piezoelectric layer by reflectors. In thin film bulk acoustic resonator (FBAR) technology, the reflector is simply air. In solid mounted resonator (SMR) technology, a Bragg reflector is used at the bottom side of the piezoelectric layer (Fig.1). The main application of FBAR and SMR technologies today is the fabrication of RF filters [2].

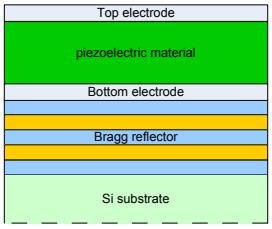


Fig. 1. Solid mounted resonator schematic cross section

In the present work, a SMR is used. An appropriate resonator model for simulation purpose is the modified Butterworth-Van Dyke equivalent circuit presented in Fig.2 [3][4]. The parameters of this model are R_m, C_m and L_m, respectively the acoustic resistor, capacitor and inductor, C_e the capacitor between the two electrodes and R_e the resistance of the electrodes. From the circuit model, we define other important parameters of the resonator as follows:

- The motional frequency:

$$f_m = 1/\left(2\pi\sqrt{L_m C_m}\right), \quad (1)$$

- The motional quality factor:

$$Q_m = 1/\left(2\pi f_m R_m C_m\right), \quad (2)$$

978-1-4244-2182-4/08/$25.00 ©2008 IEEE.

- The coupling factor which models the electro-acoustic interaction:

$$k = \sqrt{C_m/C_e}. \quad (3)$$

B. Oscillator modeling

The oscillator is decomposed into two parts for the purpose of the analysis: the active part and the resonator (Fig.2). The active part is modeled by its equivalent admittance, consisting of a negative conductance G_{osc} for compensation of resistive loss into the resonator, and of a capacitance C_{osc}. The variable capacitance C_t modeling the capacitor bank is included into the resonator part.

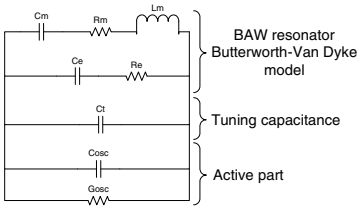


Fig. 2. Oscillator simplified model

If we denote $Y_{resonator} = G_{resonator} + j \cdot B_{resonator}$ the admittance of the resonator, and $Y_{osc} = G_{osc} + j \cdot B_{osc}$ that of the active circuit, the oscillation frequency f_{osc} is defined by:

$$B_{osc}(f_{osc}) + B_{resonator}(f_{osc}) = 0 \quad (4)$$

and the Barkhausen oscillation startup conditions at the oscillation frequency are:

$$G_{osc}(f_{osc}) + G_{resonator}(f_{osc}) < 0. \quad (5)$$

For analysis purpose, let us now define the following additional parameters:

- the effective coupling factor:

$$k' = \sqrt{C_m/(C_e + C_t + Cosc)} = \sqrt{C_m/C_p}, \quad (6)$$

- the parallel resonant frequency of the system, which is the potential oscillation frequency:

$$f'_p = f_m \times \sqrt{(C_p + C_m)/C_p} \approx f_m \times \sqrt{1 + k'^2}, \quad (7)$$

- the conductance of the oscillator at the oscillation frequency f'_p:

$$G_m = \frac{1}{R_m + R_m Q_m^2 k'^4/(1 + k'^2)}. \quad (8)$$

C. Frequency tuning

Compared to crystals used in reference frequency applications (for instance the CX-3225SB), BAW resonators have a poor frequency accuracy. The consequence is that a relatively wide tuning range is required to correct for frequency errors. From a study of the frequency dispersion of the BAW resonator, we concluded that a tuning range of about $\pm$1000ppm was necessary. As noticeable from equations (6) and (7), frequency tuning by variation of capacitance C_t has the effect of changing the effective coupling factor k'. This is illustrated by Fig.3 that shows the normalized oscillation frequency as a function of k'. When increasing capacitance C_t, both the coupling factor k' and the oscillation frequency decrease. The frequency range is bounded by two limits. First, the effective coupling factor cannot be higher than the physical coupling factor of the BAW resonator, which is determined by the process. This sets the upper limit of the tuning range. The lower limit is related to equation (8), which indicates that the oscillator conductance increases rapidly when k' decreases, i.e. when frequency decreases. At a certain point, the impedance becomes so low that startup conditions are not met anymore.

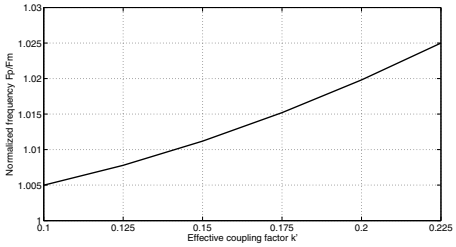


Fig. 3. Normalized frequency versus effective coupling factor

III. OSCILLATOR DESIGN

A. Topology considerations

The oscillator is designed for a resonator with a parallel resonant frequency of 2GHz, a quality factor Q_m=1000 and a coupling factor k=1.7. We selected the so-called cross-coupled NMOS differential topology in order to obtain high negative conductance and differential outputs [5]. The BAW resonator is placed between drains of the NMOS pair. Each branch of the active NMOS pair is supplied independently by a current source. So as to minimize frequency pushing and achieve high frequency stability with temperature, the supply current is stabilized relatively to supply voltage and temperature variations. In order to assess the minimum output power required for oscillator startup, we implemented a current source programmable by steps of 250μA from 250μA to 2mA. When varying the supply current, resistors at the bottom of the differential pair are switched to keep the DC common output voltage constant over the current range (Fig.4). At DC the BAW resonator and the current source show high load impedance to the cross-coupled NMOS pair. In order to prevent a latch effect in the cross-coupled pair, positive feedback in the cross-coupled pair is suppressed at low frequencies by a high-pass filter [6][7].

Current communication systems do not yet use frequency references at 2GHz. In order to measure the oscillator performances in operational conditions, we included a divider by four and a buffer. The divider has a classical D-flip-flop architecture [5]. A buffer amplifier was added following the divider for measurement purpose of the output signal into a

50Ω impedance system. Both the frequency divider and the buffer are supplied by a low drop out voltage regulator (LDO).

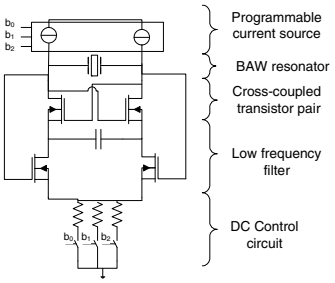


Fig. 4. Oscillator core

B. Frequency tuning

The need for frequency tuning was discussed in section II-C. A solution to implement this tuning is to realize a capacitor bank. Let us consider the bank specifications we need. The frequency uncertainty depends on four main factors:

- the initial BAW resonator error,
- the frequency drift with temperature variation,
- the frequency drift with voltage variation,
- the frequency drift with aging.

A tuning range of ±1024ppm is needed to correct these four drifts according to our calculations. The frequency step of the capacitor bank is limited by the technology. This leads us to the selection of 0.8ppm minimum frequency accuracy. To reach this small step, we decided to implement a specific fringe capacitance available in advanced CMOS technology. This takes advantage of the parasitic capacitance between two metal fingers. The main principle and model are given on Fig.5. The switched capacitance has two states:

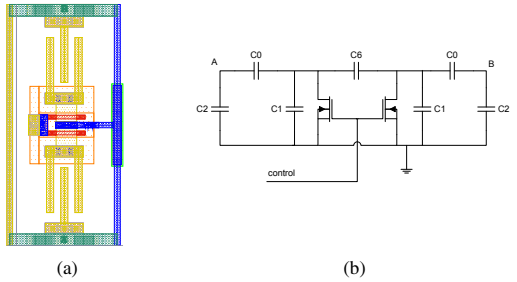
- in the off state,

$$C_{AB}^{off} = (C_0 + C_2)/2, \tag{9}$$

- in the on state,

$$C_{AB}^{on} = \left(C_2 + \frac{C_0 \cdot (2 \cdot C_6 + C_1)}{C_0 + 2 \cdot C_6 + C_1}\right)/2. \tag{10}$$

In the off state, a parasitic capacitance still remains which limits the tuning range. We have to find a trade-off between the step capacitance, $\Delta C_p = \Delta C_t = C_{AB}^{on} - C_{AB}^{off}$, and the parasitic capacitance C_{AB}^{off}.These two parameters induce the frequency step of the bank given by (7). With this in mind, we choose $C_{AB}^{on} = 1.2 \cdot C_{AB}^{off}$, corresponding to a parasitic capacitor of 310aF and a step capacitance of 64aF for a 0.8ppm accuracy. Two capacitor banks are implemented to minimize the global parasitic capacitance. The first one, called coarse bank, is dedicated to the correction of the initial error due to the BAW resonator caused by process variations. Nominally, this bank enables to correct frequency errors up to ±1024ppm. The second one, called fine bank, gives a nominal tuning range capability of ±102.4ppm. It has to cope with frequency variations that can occur while the frequency reference is running. To avoid any frequency dead zone, an overlap between the two banks is introduced as presented in Fig.6. Since the frequency step of the coarse bank should be low enough to fulfill this requirement, we choose a nominal step of 64ppm. In order to lower the differential non linearity (DNL) and the integral non linearity (INL) of the two banks, a thermometric code is implemented.



(a) (b)

Fig. 5. Fringe capacitance layout and electrical model including parasitics (C_1, C_2, C_6)

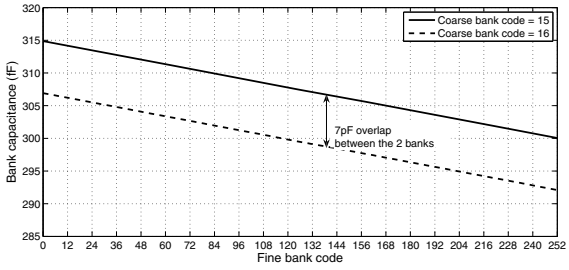


Fig. 6. Overlap between the fine and the coarse banks for coarse bank codes 15 and 16

TABLE I
SUMMARY OF THE CAPACITOR BANKS PERFORMANCES

	Frequency		ΔC_t
Coarse bank tuning range	±1024 ppm	±2 MHz	±128 fF
Coarse bank tuning step	64 ppm	132 kHz	8 fF
Fine bank tuning range	±102.4 ppm	±200 kHz	±8.16 fF
Fine bank tuning step	0.8 ppm	1.6 kHz	64 aF

C. The circuit

The complete circuit includes the oscillator core (see Fig.4), a programmable current source, two capacitor banks with their decoder and a dedicated LDO, a divider by four and a pre-buffer with a dedicated LDO, a 500MHz 50Ω matched output

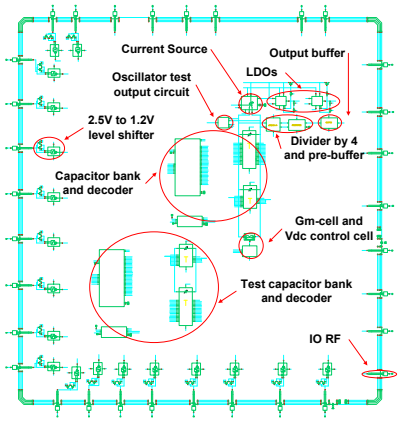


Fig. 7. Topcell schematic of the circuit

buffer and, finally, a 2GHz test output. The global circuit arrangement is presented in Fig. 7.

IV. SIMULATION RESULTS

All the blocks were individually specified and simulated. The simulations have been carried out in process corners, over voltage and temperature ranges, and took into account layout parasites. This section presents the results of the circuit. The oscillator start-up conditions (4) and (5) were checked using AC analyses. The simulation results show that the ratio between the negative conductance created by the active circuit and the BAW resonator conductance is always higher than 2 which ensure the oscillation startup. A BAW and a crystal oscillator can be compared with regard to their phase noise. A 26MHz crystal oscillator phase noise is specified to be better than -135dBc/Hz at 10kHz offset from the carrier for cellular systems. To take into account the working frequency difference between the two oscillators a correction factor has to be applied. In our case, the circuit output signal frequency is 500MHz. The correction factor is $20 \times \log(500/26) = 26dB$. Therefore, the equivalent phase noise required of the crystal oscillator would be $L(f_m) = -135+26 = -109dBc/Hz@500MHz$. The simulated phase noise of the BAW oscillator at 10kHz offset is -116dBc/Hz which is better than the -109dBc/Hz required, see Fig. 8.

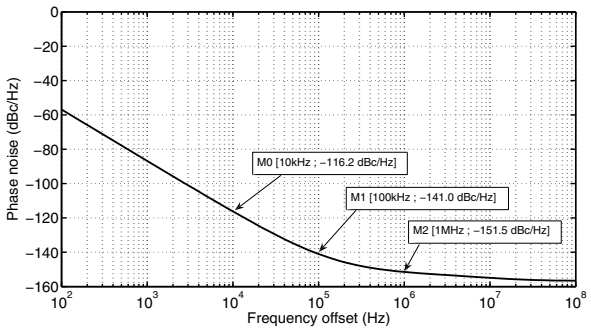


Fig. 8. Phase noise at 500MHz

The comparison with recent work given in Table II shows the relevance of our work.

TABLE II
BAW OSCILLATORS COMPARISON

Paper	type	Fosc (GHz)	Power (mW)	Phase noise at 2GHz (dBc/Hz)	Techno
[8]	Pierce (single)	1.9	0.3	-112@10kHz	BiCMOS
[1]	Butler (single)	2	4	-120@100kHz	BiCMOS
[9]	Colpitts (single)	2	6.8	-118@100kHz	SiGe BiCMOS
[6]	diff.	2.1	0.6	-122@100kHz	CMOS 0.13μm
This work	diff.	2	0.9	-128@100kHz	CMOS 65nm

V. CONCLUSION

We implemented a new digitally-tuned BAW based frequency reference in a 65nm CMOS process. A high spectral purity (-128dBc/Hz before the divider and -141dBc/Hz after at 100kHz from the carrier frequency) has been shown. Low power consumption (0.9mW) has been reached. We designed a digitally controlled capacitor bank enabling a 4MHz frequency tuning range with a 1.6kHz step using small fringe capacitors. The 200Hz step (0.1ppm) required for GSM/3G applications can be achieved with the bank control. As discussed above, these new oscillator key features bring possibilities to substitute it to crystal oscillators. This will be confirmed after full characterization of the circuit under manufacturing.

ACKNOWLEDGMENTS

The authors want to acknowledge all the ICRF engineers for their support and Frederik Vanhelmont, Andr Jansman from NXP Research for providing us with the BAW resonator.

REFERENCES

[1] F. Vanhelmont et al., "A 2 GHz Reference Oscillator incorporating a Temperature Compensated BAW Resonator," *Ultrasonics Symposium, 2006. IEEE*, 2006, pp. 333-336.
[2] H. Loebl et al., "Narrow band bulk acoustic wave filters," *Ultrasonics Symposium, 2004 IEEE*, vol. 1, 2004, pp. 411-415 Vol.1.
[3] K. Lakin, "Thin film resonators and filters," *Ultrasonics Symposium, 1999. Proceedings. 1999 IEEE*, vol. 2, 1999, pp. 895-906 vol.2.
[4] Marc-alexandre Dubois, "Thin film bulk acoustic wave resonators: a technology overview," *MEMSWAVE 03*, Jul. 2003.
[5] B. Razavi, *RF Microelectronics*, Prentice Hall PTR, 1997.
[6] S.S. Rai and B.P. Otis, "A 600 μW BAW-Tuned Quadrature VCO Using Source Degenerated Coupling," *IEEE Journal of Solid-State Circuits*, vol. 43, 2008, pp. 300-305.
[7] D. Ruffieux, "A high-stability, ultra-low-power quartz differential oscillator circuit for demanding radio applications," *28th ESSCIRC*, 2002, pp. 85-88.
[8] B. Otis and J. Rabaey, "A 300W 1.9-GHz CMOS oscillator utilizing micromachined resonators," *IEEE Journal of Solid-State Circuits*, vol. 38, 2003, pp. 1271-1274.
[9] S. Razafimandimby et al., "A 2GHz 0.25m SiGe BiCMOS Oscillator with Flip-Chip Mounted BAW Resonator," *ISSCC 2007. Dig.Tech.l Papers. IEEE International*, 2007, pp.580-623.

8.12 J.-F. Bercher and C. Berland, "Envelope and phase delays correction in an EER radio architecture," *Analog Integrated Circuits and Signal Processing*, vol. 55, pp. 21–35, Apr. 2008.

Analog Integr Circ Sig Process (2008) 55:21–35
DOI 10.1007/s10470-008-9147-z

Envelope and phase delays correction in an EER radio architecture

J.-F. Bercher · C. Berland

Received: 29 March 2007 / Revised: 10 December 2007 / Accepted: 29 January 2008 / Published online: 26 February 2008

Abstract This article deals with synchronization in the Envelope Elimination and Restoration (EER) type of transmitter architecture. To illustrate the performances of such solution, we choose to apply this architecture to a 64 carriers 16QAM modulated OFDM. We first introduce the problematic of the realisation of a highly linear transmitter. We then present the Envelope Elimination and Restoration solution and draw attention to its major weakness: a high sensitivity to desynchronization between the phase and envelope signal paths. To address this issue, we propose an adaptive synchronization algorithm relying on a feedback loop, a Least Mean Square formulation and involving an interpolation step. It enables the correction of delay mismatches and tracking of possible variations. We demonstrate that the quality of the interpolator has a direct impact on Error Vector Magnitude (EVM) value and output spectrum. Implementation details are provided along with an analysis of the behaviour and performances of the method. We present Agilent-ADS and Matlab simulation results and then focus on the enhancement of the transmitter performances using the proposed algorithm.

Keywords EER radio transmitter · Adaptive algorithm · Synchronisation · Transmitter linearisation · Interpolation

J.-F.Bercher (✉) · C. Berland
Université Paris-Est, ESYCOM-ESIEE, Cité Descartes, BP99,
93162 Noisy-le-Grand Cedex, France
e-mail: jf.bercher@esiee.fr

C. Berland
e-mail: c.berland@esiee.fr

1 Introduction

Recent radiocommunication systems aiming at high data rate are based on efficient modulation schemes in which Quadrature Amplitude Modulations (QAM) are obviously preferred to frequency or phase modulation. While the 3GPP standard employs QPSK and 16QAM, higher data rate are achieved (in WLAN system for example) using Orthogonal Frequency Data Multiplexing (OFDM) modulation. Although these modulations are suitable for signal processing, the realization of the RF front end, particularly the transmitter, becomes more and more complex.

In the conception of a transmitter, it is essential to achieve both efficiency and linearity. In radiocommunication standards, the quality of the transmitted signal is well defined, usually in terms of output spectrum, Adjacent Channel Power Ratio (ACPR) and Error Vector Magnitude (EVM). When the modulated signal presents an envelope variation, similar to an amplitude modulated signal, the compression effect of the Power Amplifier (PA) generates intermodulation products which directly impact the three mentioned figures of merit. This implies that with a classic class A PA, the input signal would have to present a mean power (depending on the type of modulation) lower than the input compression point: the difference is quantified in terms of back off. In order to gain in efficiency, a linearization system is often preferred to a linear amplification.

In this article, we deal with the *Envelope Elimination and Restoration* (EER) principle, and illustrate it in the case of a 16QAM 64 carriers OFDM modulation. The EER principle was proposed by Kahn in 1952 [14] and is based on the splitting of a modulated signal into two signals. The first one is a constant envelope phase modulated signal, while the second one is the envelope of the original signal. In the original EER transmitter, the splitting is realized in

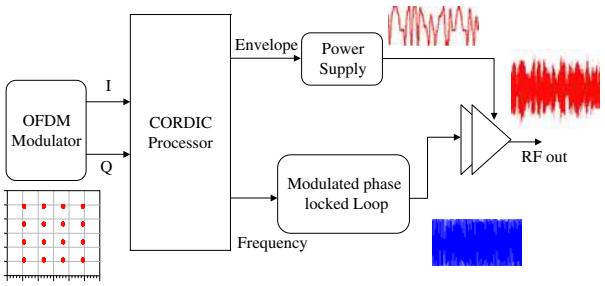


Fig. 1 Polar transmitter using a CORDIC processor

an analog way using a limiter and an envelope detector. The phase modulated signal is the input of an efficient PA whereas the envelope signal is sent to a switching power supply which feeds the last stage of the PA. The transmitter now evaluates toward a fully digital transmitter [23] as shown in Fig. 1.

The digital creation of the two signals is achieved using a CORDIC processor [22]. The principle of this processor relies on the rotation of a vector (X_i, Y_i) to a new vector. In our case, the final vector is (Z_i, θ), where Z_i is the magnitude of the vector and the rotation angle is our phase. VLSI implementation can profit from of a suitable iterative formulation. The phase signal is sent to the PA amplifier using a modulated PLL. This solution is preferred to a I/Q RF modulator in order to reduce transmitted spurious emissions. To limit distortion, the envelope signal is usually amplified with a class S modulator using a $\Sigma\Delta$ modulator rather than a classic PWM. The final amplification is realized by a high efficiency amplifier. As the input signal is a constant envelope signal, switched class of power amplifiers are preferred due to their high efficiency. As demonstrated in [20], the class E power amplifier is very suitable for this application since the expression of the output voltage is directly proportional to its power supply. This property is mandatory as we intend to reinject the envelope variation through the modulation of its power supply.

However the main drawback of this architecture is its sensitivity to delay mismatch between the two signals.

The delays introduced by the two paths can be mismatched due to pipeline differences in the paths and the delay in the anti-alias filter (amplitude path), as well as small contributions from other analog delays [21]. Furthermore, in a production environment, delays should be matched to variations in process including supply voltage, frequency, output power and temperature [18]. This usually requires careful factory calibration procedures. The mismatch deteriorates both the EVM and the output spectrum of the transmitted signal [4, 24] and has to be corrected. A linear interpolation was suggested in [11] to compensate for the mismatch, while a group delay equalizer was proposed in [17]. In these two cases, it still remains to identify the delay mismatch and track its possible variations. In this article, we propose a scheme and an associated algorithm that covers the whole calibration problem: identification, correction and tracking.

In the first part, we demonstrate the sensitivity of EER applied to a 16QAM 64 carriers OFDM modulation and bring forward the maximum tolerable delay mismatch for this modulation. The second part presents an efficient algorithm which corrects this default. We then focus on the implementation of the algorithm and on the importance of the interpolation filter used to resynchronize the signals. In the final part, an analysis of the behaviour and performances of the algorithm is provided. Simulation results performed on ADS are presented and show the performances achieved with this solution in terms of output spectrum, EVM and ACPR.

2 Impact of delay mismatches on an OFDM modulation

Using a 16QAM 64 carriers OFDM modulation is an interesting case study for the validation of this kind of architecture because of its high Peak to Average Power Ratio (PAPR). In fact, for an OFDM modulation, the PAPR is equal to the number of carriers, which corresponds to $10\log(64) = 18\,\text{dB}$ in our simulation case. A few results have already been presented on the study of the whole transmitter with this modulation [2] and this article indicates that the critical specification is the synchronization of signals. Figures 2 and 3 present the impact of delay mismatch in the range of $\pm\frac{T}{3}$, with T the time symbol, on the output spectrum and on the EVM. Table 1 gives the values of ACPR for these delays. The symbol rate is 20 MHz.

It is important to notice that the EVM calculation for OFDM modulations must be realized on each carrier separately and then averaged. Figure 4 shows that the effect of

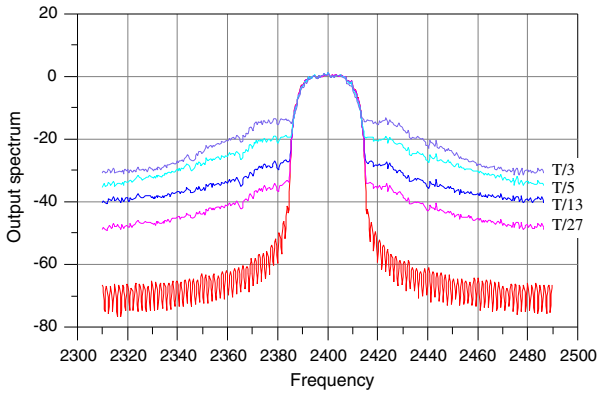


Fig. 2 Impact of the delay mismatch on output spectrum (taken in 200 kHz bandwidth)

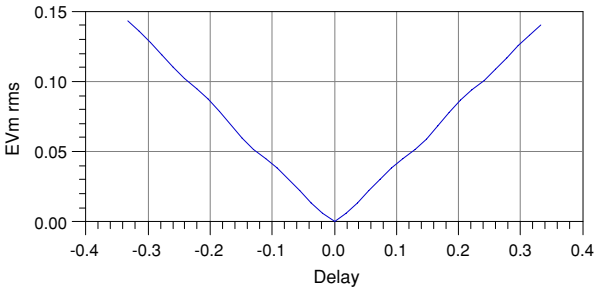


Fig. 3 Impact of the delay mismatch on the EVM

Table 1 ACPR obtained for different delay mismatches. It is evaluated for the full 64 carriers 20 MHz OFDM modulation, with a channel spacing of 25 MHz

Delay	*T*/27	*T*/13	*T*/5	*T*/3
ACPR (dB)	−37	−29.7	−20.7	−15.7

the desynchronization on the first subcarrier for delays *T*/27 and *T*/5 acts as additional noise.

As a guideline, we take specifications from 802.11.a standard: the rms EVM is specified at a maximum of 11.22% and the output spectrum has to remain under −40 dBc above a 30 MHz frequency offset from the carrier measured in a 1 MHz resolution bandwidth. When analysing the EVM results, we can observe that a desynchronization of *T*/3 produces an EVM value of 14% and when considering the output spectrum, in 1 MHz bandwidth, the value at 30 MHz from the carrier is about −7.5 dBc. In fact, the value of −40 dBc is only obtained for a delay of *T*/27 which gives an EVM value of about 0.6% This demonstrates that the synchronization of signals impacts so strongly on the output spectrum that it is mandatory to implement a correction algorithm.

3 The correction algorithm

This article is an extended version of [3] where we presented a preliminary version of the algorithm. In this version, the algorithm is presented in more details and implementation issues are discussed. More specifically, we discuss the role, performances and implementation of the interpolation step and provide a detailed analysis of the algorithm performances.

3.1 Envelope and phase alignment

According to the central limit theorem, for complex modulation schemes such as the OFDM, when the number of subcarriers is large, the emitted signal can be approximated as a gaussian distributed complex random variable. For a narrowband stationary signal written as

$$\xi(t) = x(t)\cos(\omega_0 t) + y(t)\sin(\omega_0 t) \tag{1}$$

it is well known that $R_{xx}(\tau) = R_{xy}(\tau)$ and that $R_{xy}(\tau) = -R_{yx}(\tau)$, where R_{xx} and R_{xy} are respectively the autocorrelation and intercorrelation functions. The processes $x(t)$ and $y(t)$ are always uncorrelated at the same instant, that is $R_{xy}(0) = 0$. If $\xi(t)$ is normal, then $x(t)$ and $y(t)$ are independent at the same instant. The complex gaussian process $\xi(t)$ is completely characterized by its mean and autocorrelation function

$$R_{\xi\xi}(\tau) = R_{xx}(\tau)\cos(\omega_0\tau) + R_{xy}(\tau)\sin(\omega_0\tau). \tag{2}$$

Let us denote $S_{\xi\xi}(f)$ and $S_{xx}(f)$ the spectra of $\xi(t)$ and $x(t)$, that is the Fourier transforms of the autocorrelation functions $R_{\xi\xi}(\tau)$ and $R_{xx}(\tau)$. When $S_{\xi\xi}(f)$ is symmetric with central frequency f_0, the in-phase and quadrature components $x(t)$ and $y(t)$ are uncorrelated, that is $R_{xy}(\tau) = 0$. The baseband spectrum $S_{xx}(f)$ is generally proportional to the square of the transfer function of the emission filter, which shall be shaped as a square-root Nyquist filter. Consequently, the autocorrelation function $R_{xx}(\tau) = R_{yy}(\tau)$, which is the inverse Fourier transform of the baseband spectrum, is the impulse response h of a (full) Nyquist filter. The process $\xi(t)$ can also be written as

$$\xi(t) = \rho(t)\cos(\omega_0 t - \phi(t)) \tag{3}$$

where $\rho(t)$ is the envelope and $\phi(t)$ is the phase process, with

$$\begin{cases} x(t) = \rho(t)\cos(\phi(t)), \\ y(t) = \rho(t)\sin(\phi(t)). \end{cases}$$

In the case of a complex gaussian process, it is well known that the envelope and phase are independent *at the same instant* and respectively distributed according to the

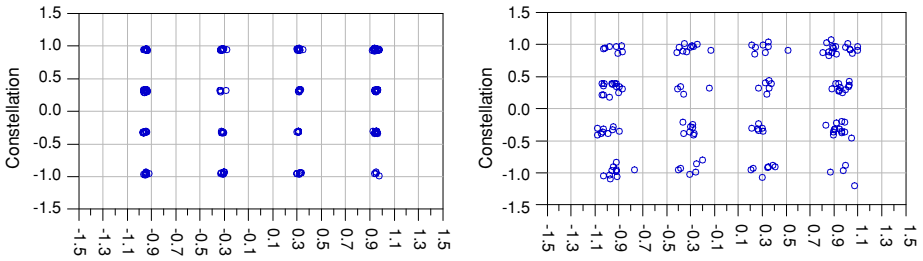


Fig. 4 Impact of the delay mismatch on the constellation for T/27 and T/5

Rayleigh and uniform distributions. The case of delayed envelope and phase is less known. In fact, it appears that for gaussian processes, the envelope $\rho(t)$ and phase $\phi(t-\Delta)$ are also Rayleigh and uniform distributed, and are *always* independent (see [4]) with no reference to the correlation coefficient, whatever the delay between envelope and phase components.

In consequence, the output do not convey any information on the time alignment or mismatch between the envelope and phase components. As a result, it is *not* possible to correct the relative delay between the envelope and phase components from the sole observation of the system output. Calibration then needs to rely on a feedback loop and involve a direct comparison between the initial "aligned" signal and the observed signal.

In a calibration step without data transmission, we can also modify the modulation scheme and design non-gaussian sequences with some dependence between envelope and phase. In such a situation, a 'contrast' based on the output properties may be devised in order to align the components. However, we will focus here on the feedback solution that preserves the data and modulation technique.

3.2 The compensation algorithm

Let us denote $z(t)$ the output of the system. Due to the delays Δ_1 and Δ_2 that affect the envelope and phase components, we have

$$\begin{aligned} z(t) &= \rho(t-\Delta_1)\cos(\omega_0 t - \phi(t-\Delta_2)) \\ &= \rho(t-\Delta_1)\cos(\phi(t-\Delta_2))\cos(\omega_0 t) \\ &\quad + \rho(t-\Delta_1)\sin(\phi(t-\Delta_2))\sin(\omega_0 t). \end{aligned}$$

We propose here to correct the delays using an adaptive precompensation. The synchronization algorithm relies on the idea of introducing two advances μ_1 and μ_2 in order to precompensate the delays, as illustrated in Fig. 5. In such a case, the output becomes

$$z(t) = \rho(t+\mu_1-\Delta_1)\cos(\omega_0 t - \phi(t+\mu_2-\Delta_2)) \tag{4}$$

and we will adjust the advances μ_1 and μ_2 in order to minimize a statistical distance between $z(t)$ and $\xi(t)$. A natural criterion is the minimization of the quadratic distance

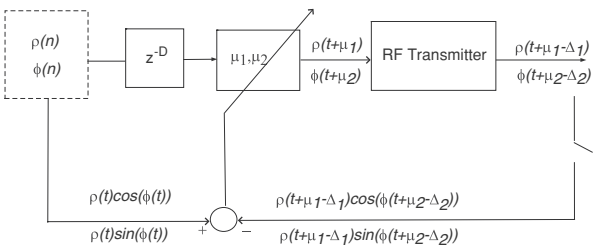


Fig. 5 Principle of delays correction

$$J(\mu_1,\mu_2;t) = \mathrm{E}\left[|\xi(t)-z(t)|^2\right], \tag{5}$$

where $\mathrm{E}[\bullet]$ is the statistical expectation operator. This expression can also be rewritten as

$$\begin{aligned} J(\mu_1,\mu_2;t) = &\mathrm{E}\left[e_x(t)^2\right]\cos(\omega_0 t)^2 + \mathrm{E}\left[e_y(t)^2\right]\sin(\omega_0 t)^2 \\ &+ \mathrm{E}\left[e_x(t)e_y(t)\right]\sin(2\omega_0 t) \end{aligned}$$

where $e_x(t)$ and $e_y(t)$ are the errors for the in-phase and quadrature components respectively,

$$e_x(t) = (\rho(t)\cos\phi(t) - \rho(t_1)\cos\phi(t_2)) \tag{6a}$$

$$e_y(t) = (\rho(t)\sin\phi(t) - \rho(t_1)\sin\phi(t_2)) \tag{6b}$$

and where, in order to simplify the expressions, we noted

$$t_1 = t+\mu_1-\Delta_1 \text{ and } t_2 = t+\mu_2-\Delta_2. \tag{7}$$

After time averaging, this simply reduces to

$$J(\mu_1,\mu_2) = \frac{1}{2}\left\{J_x(\mu_1,\mu_2) + J_y(\mu_1,\mu_2)\right\} \tag{8}$$

with

$$J_x(\mu_1,\mu_2) = \mathrm{E}\left[e_x(t)^2\right] \text{ and } J_y(\mu_1,\mu_2) = \mathrm{E}\left[e_y(t)^2\right]. \tag{9}$$

The global criterion equals the sum of two elementary criteria on the in-phase and quadrature components. We can readily obtain the same criterion (up to a factor) using the demodulated, baseband, version of the signal.

In practice, we indeed work in baseband with digital signals. After sampling, we compare the digital input signal to the sampled baseband output. In the following, we will keep T for the symbol period and note T_s the sampling period. Consequently, we will note the different discrete time indexes as follow

$$\begin{cases} t(n) = nT_s \\ t_1(n) = nT_s + \mu_1 - \Delta_1 \\ t_2(n) = nT_s + \mu_2 - \Delta_2. \end{cases} \tag{10}$$

The output sampling clock does not need to be synchronous to the input: it may be a divided version of the input clock, and any propagation delay will be absorbed in the correction procedure.

Since it is simpler to generate delayed signals than advanced signals, we introduce a small processing delay D in Fig. 5.

Let us consider the criterion $J_x(\mu_1,\mu_2)$ in (9) on the in-phase component. Developing and taking into account the independence between $\rho(t_1)$ and $\phi(t_2)$ lead to

$$J_x(\mu_1,\mu_2) = 2C_x(0,0) - 2C_x(\mu_1-\Delta_1,\mu_2-\Delta_2) \tag{11}$$

where

$$C_x(\tau_1,\tau_2) = \mathrm{E}[\rho(t)\cos(\phi(t))\rho(t-\tau_1)\cos(\phi(t-\tau_2))] \tag{12}$$

Analog Integr Circ Sig Process (2008) 55:21–35

with $\tau = 1 = \mu_1 - \Delta_1$ and $\tau_2 = \mu_2 - \Delta_2$, is a kind of 'correlation function'. The term $C_x(0,0)$ simply reduces to $C_x(0,0) = \mathrm{E}\left[\rho(t)^2\right]\mathrm{E}\left[\cos(\phi(t))^2\right] = R_{xx}(0)$, the variance of the in-phase $x(t)$ component. It is important to note that $C_x(\tau,\tau)$, obtained with $\tau_1 = \tau_2 = \tau$ is nothing else but the correlation function $R_{xx}(\tau)$. Since we know that the correlation function is proportional to the shaping filter h, it appears that the behaviour of the criterion $J_x(\mu_1,\mu_2)$ is closely related to the shaping filter. Regarding the quadrature component and criterion $J_y(\mu_1,\mu_2)$, the same conclusions and formulas found in (11, 12) are easily obtained by substituting x by y and cos by sin.

In the case of an OFDM modulation, the criterion $J(\tau_1,\tau_2)$ was evaluated numerically by Monte Carlo simulations with a square root Nyquist filter (square root raised cosine with 0.5 roll-off). This is illustrated in Fig. 6 for delays (advances) between $-4T$ and $4T$. We can recognize here the general shape of the (inverted) impulse response of a raised cosine, somewhat distorted and modulated. The criterion does not only present a global minimum at $\tau_1 = 0$, $\tau_2 = 0$, but also several other minima. Derivation of a closed-form formula for criteria (9) involving (12) is a challenging if not impossible task. However, in Fig. 7, where the criterion for delays less than $1.5T$ is shown, we clearly observe that any descent algorithm will avoid local minima for delays $\Delta_1, \Delta_2 \leq T$, with initial conditions set to zero.

3.3 Gradient algorithm

Since we do not have a closed-form for the criterion nor a direct explicit solution for its global minimizer, we need to exhibit the solution using a descent algorithm. We simply

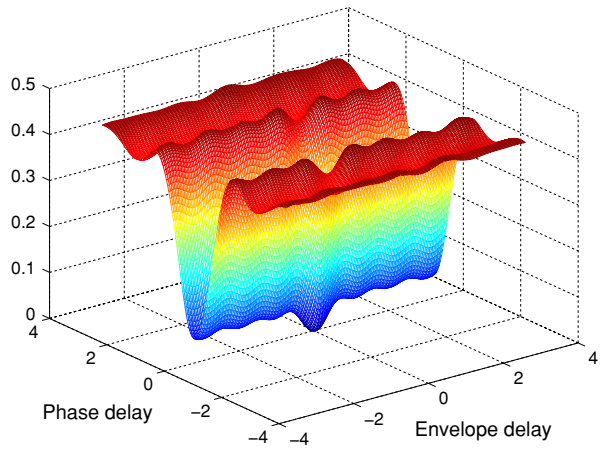


Fig. 6 Shape of the criterion for the envelope and phase delays $\in [-4T, 4T]$, with T the symbol period. The shape of the criterion is related to the impulse response of the shaping filter and presents local minima

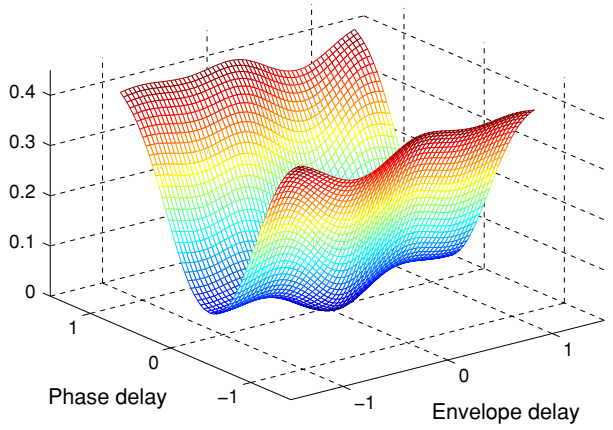


Fig. 7 Shape of the criterion for the envelope and phase delays $\in [-3T/2, 3T/2]$, with T the symbol period

use a gradient algorithm that consists in iterating the following formulas:

$$\begin{cases} \mu_1(n+1) = \mu_1(n) - \gamma_1(n)\frac{\partial J(\mu_1,\mu_2)}{\partial \mu_1}\Big|_{\mu_1=\mu_1(n)} \\ \mu_2(n+1) = \mu_2(n) - \gamma_2(n)\frac{\partial J(\mu_1,\mu_2)}{\partial \mu_2}\Big|_{\mu_2=\mu_2(n)} \end{cases} \tag{13}$$

where $\gamma_1(n)$ and $\gamma_2(n)$ are two adaptation steps that possibly depend on the iteration index n. The gradients are given by

$$\frac{\partial J(\mu_1,\mu_2)}{\partial \bullet} = \frac{1}{2}\left[\frac{\partial J_x(\mu_1,\mu_2)}{\partial \bullet} + \frac{\partial J_y(\mu_1,\mu_2)}{\partial \bullet}\right] \tag{14}$$

and we readily obtain

$$\frac{\partial J_x(\mu_1,\mu_2)}{\partial \mu_1} = -E\left[\frac{d\rho(u)}{du}\Big|_{u=t_1(n)} \cos\phi(t_2(n))e_x(t(n))\right] \tag{15a}$$

$$\frac{\partial J_y(\mu_1,\mu_2)}{\partial \mu_1} = -E\left[\frac{d\rho(u)}{du}\Big|_{u=t_1(n)} \sin\phi(t_2(n))e_y(t(n))\right] \tag{15b}$$

and

$$\frac{\partial J_x(\mu_1,\mu_2)}{\partial \mu_2} = -E\left[\frac{d\cos\phi(u)}{du}\Big|_{u=t_2(n)} \rho(t_1(n))e_x(t(n))\right] \tag{16a}$$

$$\frac{\partial J_y(\mu_1,\mu_2)}{\partial \mu_2} = -E\left[\frac{d\sin\phi(u)}{du}\Big|_{u=t_2(n)} \rho(t_1(n))e_y(t(n))\right] \tag{16b}$$

with e_x and e_y defined by (6a, 6b).

The update equations are obtained using these gradients in (13). However, we do not have the analytical expressions of the statistical expectations involved in these

formulas. Therefore, we have to resort to use a stochastic approximation of these theoretical recursions. A popular solution in adaptive filtering is the *Least Mean Squares* (LMS) algorithm that simply consists in omitting the statistical expectation. The LMS then involves the *instantaneous gradient* rather than the (correct) statistical average. Furthermore, the equations are updated at each new sample. This gives

$$\mu_1(n+1) = \mu_1(n) + \gamma_1(n)\left.\frac{d\rho(u)}{du}\right|_{t_1(n)} \times \left(\cos\left(\phi(t_2(n))\right) e_x(t) + \sin\phi(t_2(n)) e_y(t(n))\right) \quad (17)$$

$$\mu_2(n+1) = \mu_2(n) + \gamma_2(n)\rho(t_1(n)) \times \left(\left.\frac{d\cos\phi(u)}{du}\right|_{t_2(n)} e_x(t) + \left.\frac{d\sin\phi(u)}{du}\right|_{t_2(n)} e_y(t(n))\right). \quad (18)$$

Practical implementation of formulas (17) and (18) requires

- computation of the errors $e_x(t(n))$ and $e_y(t(n))$ defined in (6a) and (6b) which results from the comparison of the system input and output,
- computation of the derivatives that can be simply approximated by finite differences of $\rho(t_1(n))$ and $\cos(\phi(t_2(n)))$.

Of course, the algorithm can be simplified by considering a sole error component rather than two. For instance, one can simply put $e_y = 0$ in the previous equations. In practice, simulations show that the gain associated with the second component is extremely small.

For formulas (17) and (18), the computational load is about 8 real multiplications per iteration. However, since the derivatives must be computed at time $t_1(n)$ and $t_2(n)$, that is at the output of the corrected system, $\rho(t)$ and $\cos(\phi(t))$ must be separately accessible. This implies a quadrature demodulation before the feedback loop. Furthermore, with this approach, we need to adjust two advances μ_1, μ_2 and apply them to the input signal. Adopting very high sampling frequencies in order to get the required precision may not be an efficient solution. Digital interpolation is a more effective solution as it keeps reasonable sampling frequencies and save consumption.

4 Impact of the interpolator

Contrary to what we indicated in [3], the interpolation procedure has a significant impact on the performances of the algorithm in terms of the EVM and output spectrum.

According to the Shannon–Nyquist sampling theorem, we know [29] that any band-limited signal $x(t)$ can be recovered exactly from its samples $x(m) = x(mT_s)$ taken at the sampling frequency $1/T_s$ by the formula

$$x(t) = \sum_m x(m)\mathrm{sinc}(\pi(t - mT_s)/T_s), \quad (19)$$

where sinc is the cardinal sine. This indicates that, in principle, the samples convey enough information to reconstruct the original signal at any desired time. In particular, it is possible to reconstruct $x(t-\tau)$, for any τ, and therefore new shifted samples $x(kT_s - \tau)$ from the original samples $x(kT_s)$, according to

$$x(kT_s - \tau) = \sum_m x(mT_s)\mathrm{sinc}(\pi(kT_s - \tau - mT_s)/T_s). \quad (20)$$

The above expression is in the form of a digital convolution and can be implemented as a filtering operation. However, because the underlying filter has an infinite (sinc) impulse response, and is non causal, practical implementation introduces truncation and delay. Another possibility is to use a convenient approximation of the ideal interpolator mentioned above. The MMSE FIR interpolator [19] is the minimum mean square error approximation of the ideal filter with finite impulse response. Although optimum implementations of these interpolators exist [6], the coefficients shall be pre-computed and tabulated for each possible fractional delay and these types of structures should be reserved for the interpolation with fixed delays.

Another class of interpolators relying on polynomial approximation can be used instead. Indeed, the Weierstrass approximation theorem states that every continuous function defined on an interval can be uniformly approximated as closely as desired by a polynomial function. We can then use a polynomial to approximate the value of the function, given a series of samples, at the desired delay. Such interpolators are especially interesting for our application since they can be described by FIR filters. They can be implemented very efficiently in hardware, and their coefficients can be computed in real time rather than taken from a table. An efficient structure was devised by Farrow [8, 15] and improvements to the structure can be found in [25, 7, 5]. In this structure, the delay is directly adjustable without modification so that it is suitable for our adaptive synchronization problem.

A related problem is Sample Rate Conversion (SRC) which is often considered in digital front ends [10, 9]. In SRC, a digital signal has to be converted into another digital signal but with a different sampling frequency. Caution must be exercised to avoid aliasing in the operation. In this situation, polynomial interpolators are usually disqualified because they do not provide enough anti-aliasing. In the last decade, following [27], many solutions

Analog Integr Circ Sig Process (2008) 55:21–35

have been developed for the synthesis of adjustable fractional delay filters with larger bands and better anti-aliasing capabilities [26, 13, 12, 30, 32]

It is worth mentioning that adjustable fractional delay filters can also be obtained using programmable allpass Infinite Impulse Response (IIR) filters [16, 31]. However, such filters are more sensitive to quantization, transients may occur when changing coefficients, and synthesis is complicated by the stability issues.

In our application, interpolation operates directly on the digital input signal, without any rate change. Aliasing can still occur due to the sampling operation of the output, which is needed for our feedback loop. However, the system specifications, in particular the power limitation in adjacent channels, severely constrain the design and limit the images of the original band-limited spectrum. Furthermore, since the over-sampling ratio T/T_s is typically greater than 5, an anti-aliasing filter can be easily designed.

The over-sampling ratio being high is an important factor since it allows the use of very low-order interpolators. Indeed, the frequency response of the corresponding filters is almost flat in magnitude and linear in phase in the band of interest. This is illustrated in Sect. 4.1. Then, in Sect. 4.2, we examine and compare the performances of the different interpolators in terms of interpolation error and EVM at the output of the transmitter.

4.1 The interpolators and their frequency responses

In this section, we choose to compare four interpolators [1, Chapter 25]: Linear, Bessel, 3rd and 5th order Lagrange interpolators. We first give the expressions of these interpolators and then compare their frequency responses.

The first interpolator, the forward linear interpolator, is the simplest, and is given by:

$$x(m,\tau) = x(m) + \tau(x(m+1) - x(m)) \tag{21}$$

where $x(m,\tau)$ represents the interpolated value of the input value at the time $(m+\tau)T_s$.

The second interpolator studied is the 3rd order Lagrange interpolator:

$$\begin{aligned} x(m,\tau) =& x(m) + \tau(x(m) - x(m-1)) \\ &+ \frac{\tau(\tau+1)}{2}(x(m+1) - 2x(m) + x(m-1)). \end{aligned} \tag{22}$$

This interpolation uses three successive points. The first part of the expression is similar to the linear interpolator and the second part is a correction term calculated using the point before and after the central one.

The third interpolator we looked at is the Bessel central difference interpolator, 4th order, described as follow:

$$\begin{aligned} x(m,\tau) =& x(m) + \tau(x(m+1) - x(m)) \\ &+ \frac{\tau(\tau-1)}{4}(x(m+2) - x(m+1) - x(m) \\ &+ x(m-1)). \end{aligned}$$

This interpolation uses four successive points and is also similar to the linear interpolation with an additional correction term. Compared to the 3rd order Lagrange formulation, the correction term is not symmetrical.

The last interpolator is the 5th order Lagrange interpolator:

$$\begin{aligned} x(m,\tau) =& \frac{(\tau^2-1)(\tau-2)\tau}{24}x(m-2) \\ &- \frac{(\tau^2-4)(\tau-1)\tau}{6}x(m-1) + \frac{(\tau^2-1)(\tau^2-4)}{4}x(m) \\ &- \frac{(\tau^2-4)(\tau+1)\tau}{6}x(m+1) \\ &+ \frac{(\tau^2-1)(\tau+2)\tau}{24}x(m+2). \end{aligned} \tag{23}$$

These different interpolators can be clearly viewed as FIR filters, where impulse responses can be deduced from the above equations. Therefore, it is certainly interesting to compare their frequency responses to the frequency response of a pure delay. This comparison is shown in Fig. 8 for the magnitude and in Fig. 9 for the phase. With the exception of the linear interpolator, it appears that the interpolators have interesting performances for normalized frequencies below 0.2 (over-sampling ratio greater than 5). While the two Lagrange interpolators show the flattest magnitude, the Bessel interpolator exhibits a better phase linearity. This information will be completed by other measures of performances, namely interpolation error and EVM at the output of the transmitter.

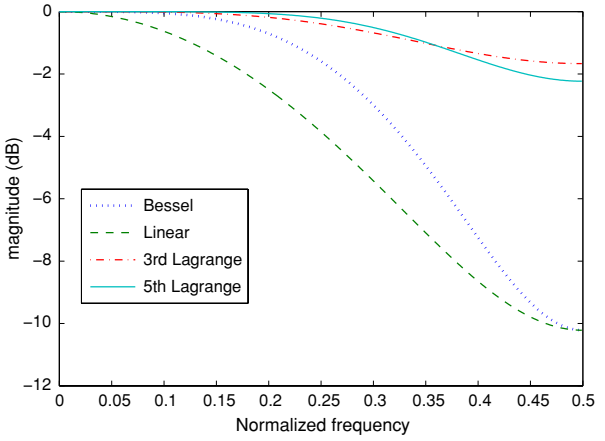


Fig. 8 Magnitudes of the frequency responses of the four interpolators

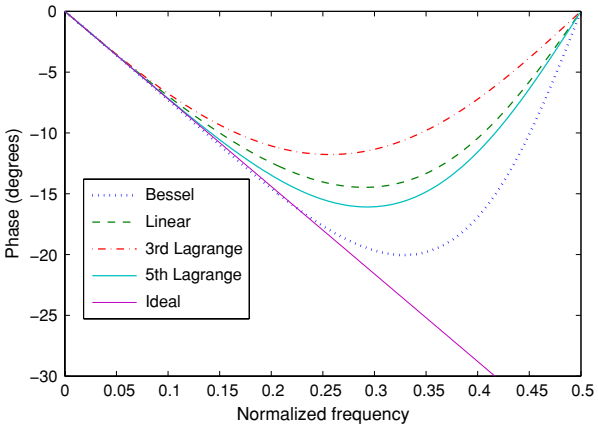


Fig. 9 Phases of the frequency responses of the four interpolators

4.2 Comparison of performances

In order to evaluate interpolation errors, we compare an original signal to a reconstructed one. This comparison is realized on the envelope of the OFDM modulated signal previously introduced. Starting with an original signal with 300 samples per time symbol, we subsample it by a factor of 100. This gives an over-sampling ratio of 3. The signal is then reconstructed by interpolating the missing 200 values between $x(m-1)$ and $x(m+1)$, and the interpolated values at a delay Δ from the sample, are compared to the value of the original signal. Performances of the different interpolators are presented in Fig. 10 in terms of the rms quadratic error, given in percent. These experimental results are confirmed by the theoretical analysis in Sect. 4.4.

Let us note T_s the sampling period and call E_o the original signal, E_i the interpolated signal. The quadratic error (rms) is expressed as:

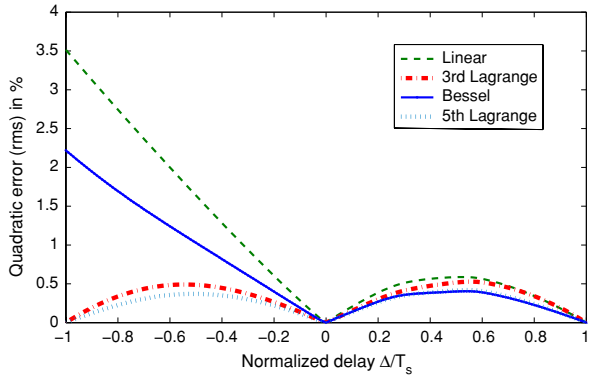


Fig. 10 Comparison of the normalized quadratic error (rms) between the four interpolators

$$Err(\Delta) = 100\sqrt{\frac{\sum_k \left(E_o(kT_s-\Delta) - E_i(kT_s-\Delta)\right)^2}{\sum_k E_o^2(kT_s-\Delta)}} \qquad (24)$$

When the interpolators are used between $x(m)$ and $x(m+1)$, their rms quadratic errors are similar with a maximum error in the middle of the two samples. The Bessel and 5th order Lagrange interpolators achieve similar results with a maximum error of about 0.4%, while the two other interpolators present less than 0.6%. However, a problem arises when the interpolator is used in the range of $x(m-1)$ and $x(m)$. The linear interpolator is obviously not up to par with a 3.5% of error followed by the Bessel with 2.5%. The asymmetry can degrade the quadratic error performances. During iterations, positive as well as negative values of τ may indeed appear. The best solution remains certainly the two Lagrange formulas which are quasi symmetrical and more appropriate to our problem. The order of the interpolator can also make a difference in terms of the complexity of the implementation.

Using ADS, the interpolators can be validated with the study of the EVM (computed for the OFDM after FFT demodulation) and the output spectrum. The simulation is realized differently than with the previous quadratic error evaluation. Here the signal sampled at $T/3$ is delayed or advanced by τ taken between ± 1 and resynchronized using the interpolator. This is a complementary analysis to the previous analysis and is better suited to standard transmitter analysis.

Figure 11 shows the EVM at the output of the transmitter and presents the same profile as the error quadratic curves in Fig. 10.

The main problem is the relatively high values of the EVM. Values reach a maximum of 2% for the 5th order Lagrange interpolator and 3.5% for the 3rd order Lagrange interpolator. The latter is not acceptable for our application. We then need to sample the signal to a higher rate.

Figure 12 presents the EVM values when changing the sampling period of the signal from $T/3$ to $T/9$ (odd values are here preferred to facilitate the demodulation of the OFDM signal in simulation). The x axis of the curve is

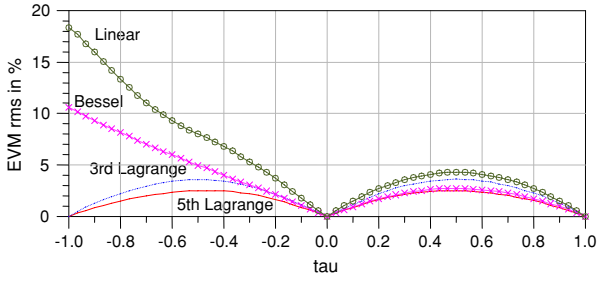


Fig. 11 Comparison of EVM values for the different interpolators

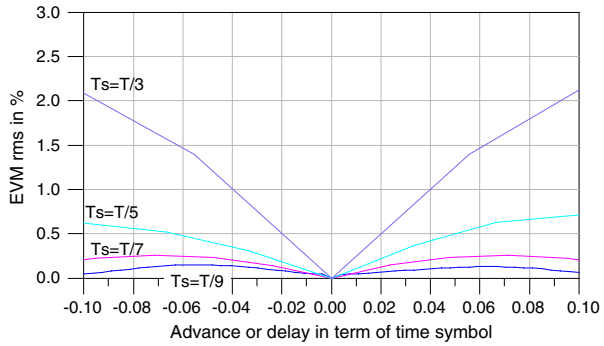


Fig. 12 EVM values for the 5th order Lagrange interpolator with different sampling periods

normalized to symbol duration. The difference between samplings at *T*/3 and *T*/9 is dramatic, going from 2% to 0.1%.

The sampling rate not only has an impact on the EVM value, it also affects the output spectrum.

4.3 Implementation using the Farrow structure

A polynomial interpolator can efficiently be implemented using a Farrow structure [8] or its extensions [27, 13, 5]. We only describe here the original Farrow structure but other possibilities will deserve further investigation. The Farrow structure relies on the parallelization of *N* FIR filters which can be expressed as

$$y_i(m) = \sum_{k=-M}^{M-1} h_i(k)x(m-k). \tag{25}$$

The output of these filters are combined so that the output signal can be expressed as

$$y(m,\tau) = \sum_{i=0}^{N-1} \tau^i y_i(m) = \sum_{i=0}^{N-1} \tau^i \sum_{k=-M}^{M-1} h_i(k)x(m-k) \tag{26}$$

For instance, rewriting the equations of the 5th order Lagrange interpolator (23) gives

$$\begin{aligned}
y_0(m) &= \frac{5}{4}x(m),\\
y_1(m) &= -\frac{2}{24}(m+2) + \frac{4}{6}x(m+1)\\
&\quad -\frac{4}{6}x(m-1) - \frac{2}{24}(m-2)\\
y_2(m) &= -\frac{1}{24}(m+2) + \frac{4}{6}x(m+1) - \frac{5}{4}x(m)\\
&\quad +\frac{4}{6}x(m-1) - \frac{1}{24}(m-2)\\
y_3(m) &= \frac{2}{24}(m+2) - \frac{1}{6}x(m+1) + \frac{1}{6}x(m-1) - \frac{2}{24}(m-2)\\
y_4(m) &= \frac{1}{24}(m+2) - \frac{1}{6}x(m+1) + \frac{1}{4}x(m) - \frac{1}{6}x(m-1)\\
&\quad +\frac{1}{24}(m-2)
\end{aligned}$$

and can be implemented on ADS as presented in Fig. 13.

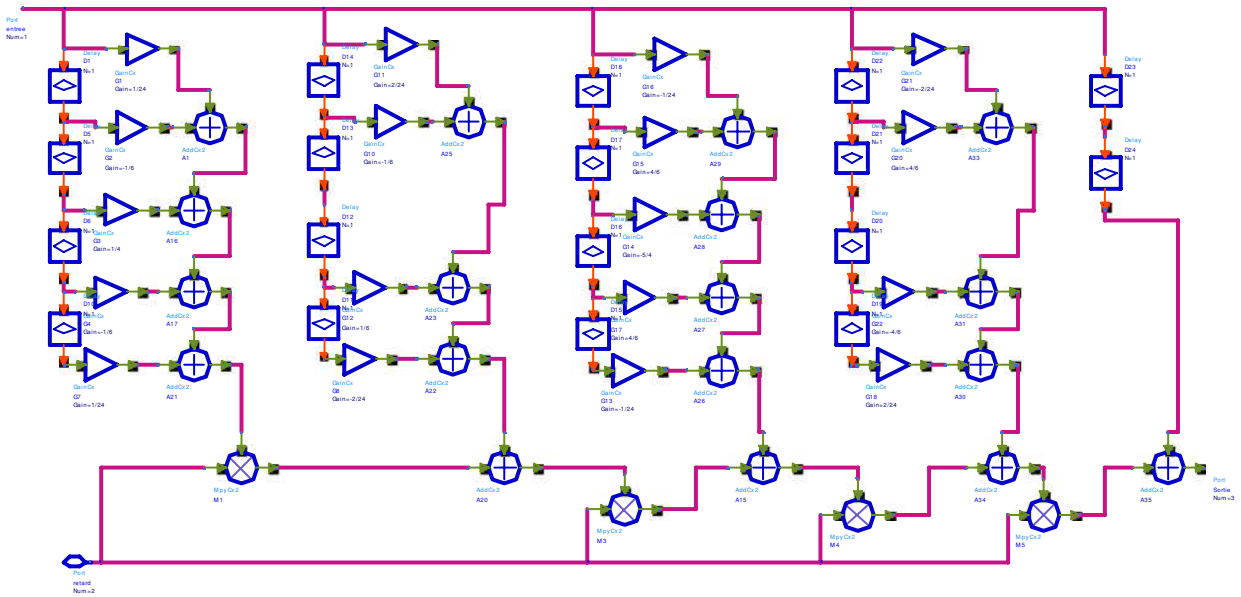


Fig. 13 Farrow structure for the 5th order Lagrange interpolator

For the 5th order Lagrange filter, this implementation uses 5 multiplications, 20 additions and 15 arithmetic divisions. As for the 3rd order Lagrange interpolator, the implementation leads to only 2 multiplications, 5 additions and simpler divisions that only requires to divide by 2. This can certainly be an argument for the choice of the interpolator, not only in terms of performances but also in terms of simplicity of implementation.

4.4 On the interpolation noise

If the interpolator were perfect, the algorithm, in the absence of observation noise, would find the *exact* solution. However, as illustrated in Figs. 10, 11 and 12, the system is characterized by an internal "self-noise" associated with the interpolator. Indeed, for equally spaced points t_0, t_1,…, t_n, the remainder of the Lagrange interpolation polynomial of degree n for estimating $x(t_0+\tau T)$ is [28]

$$n_i(t_0+\tau T) = T^{n+1}\omega(\tau)\frac{x^{(n+1)}(\xi)}{(n+1)!} \tag{27}$$

with $\omega(\tau)=\tau(\tau-1),\ldots,(\tau-n)$, and where ξ is an unknown point in the interval $[t_0,t_n]$. Thus, we can evaluate the variance $\sigma_i^2(\tau)=\mathrm{E}\left[n_i(t_0+\tau T)^2\right]$ of this interpolation noise $n_i(t)$ and obtain

$$\mathrm{E}\left[n_i(t_0+\tau T)^2\right] = (-1)^{n+1}\frac{T^{2(n+1)}}{(n+1)!}|\omega(\tau)|^2 R_{xx}^{(2(n+1))}(0) \tag{28}$$

using the relation $\mathrm{E}\left[|x^{(n+1)}(t)|^2\right]=(-1)^{n+1}R_{xx}^{(2(n+1))}(0)$.

As a result, the variance of the interpolation noise is non-stationary since it depends on the interpolation point $t_0+\tau T$. With this formula, it also appears that the higher the sampling rate and the higher the interpolation order, the lower the variance. To illustrate this, Fig. 14 compares the theoretical formula (28) to simulation results (see also the results in Fig. 10).

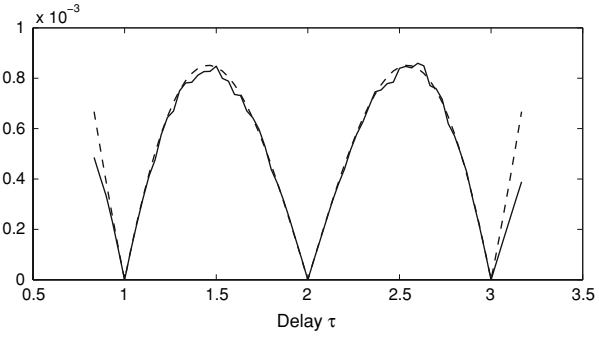


Fig. 14 Standard deviation of the interpolation noise for a Lagrange 5th order interpolator. Theoretical formula (28) (plain line) is compared to simulation results (dashed line)

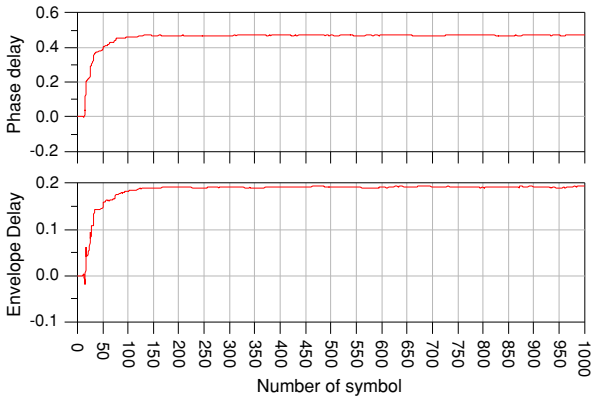


Fig. 15 Convergence of μ_1 and μ_2 to the true delays $\Delta_1=0.47T$ and $\Delta_2=0.18T$

5 Analysis of the algorithm and results

Typical results for delays $0.18T$ and $0.47T$ for envelope and phase respectively, are shown in Fig. 15. The measured EVM after convergence is only 0.7% compared to 14% without correction. The resulting spectrum is reported in Fig. 16. It shows very interesting performances: the spectrum is improved by 30 dB compared to the uncorrected case. The remaining noise floor at −50 dBc in 200 kHz bandwidth corresponds to the interpolation errors. Surprisingly, it appears that the spectrum obtained with the values at the output of the correction algorithm is slightly better than the spectrum obtained with the true values of delays. This is commented in §5.2.

However interesting these results are, it is important to examine the role of the interpolator, the impact of the values of the adaptation step and to study figures of merit such as the settling time (convergence speed), the bias and variance of results as well as the overall EVM. We evaluated these different points by Monte Carlo experiments

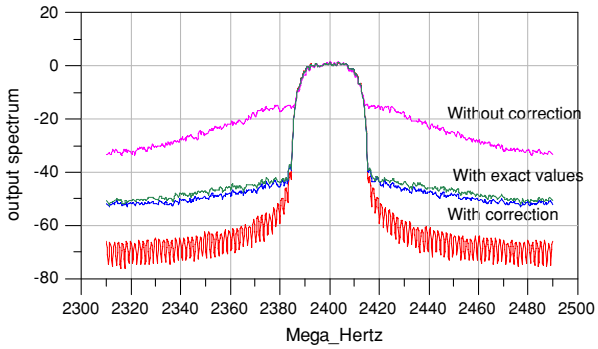


Fig. 16 Comparison of ideal, uncorrected and corrected spectra, in the case of delays $\Delta_1=0.47T$ and $\Delta_2=0.18T$

Analog Integr Circ Sig Process (2008) 55:21–35

for a varying adaptation step γ and for all delays less than T. In the study, we took the same adaptation $\gamma = \gamma_1 = \gamma_2$ step for both adaptations. Furthermore, the signal power was normalized to one in order to be independent of signal scales.

To analyse these different characteristics, we use a simplified 'toy' model that is simpler to analyse than our original problem. Let $z(t) = x(t - \Delta)$ be an observed, delayed version of an original signal $x(t)$. In order to identify and correct the delay, we adopt the following recursion

$$\mu(n+1) = \mu(n) + \gamma \left.\frac{dx(u)}{du}\right|_{u=t-\Delta+\mu(n)} [x(t) - \tilde{x}(t - \Delta + \mu(n))] \tag{29}$$

with $\tilde{x}(t - \Delta + \mu(n))$ the interpolated value at time $(t - \Delta + \mu(n))$.

5.1 Settling time

The convergence speed is an important issue for the practical use of such algorithm. With $\Delta\mu(n) = (\mu(n) - \Delta)$ small, we have

$$x(t) - \tilde{x}(t + \Delta\mu(n)) \simeq -\Delta\mu(n)\dot{x}(t) + n_i(t), \tag{30}$$

where $\dot{x}(t)$ is the derivative of process $x(t)$ and where $n_i(t)$ represents the interpolation noise. Then,

$$\begin{aligned}\Delta\mu(n+1) =& \Delta\mu(n) - \gamma\Delta\mu(n)\dot{x}(t + \Delta\mu(n))\dot{x}(t) \\ &+ \gamma\dot{x}(t + \Delta\mu(n))n_i(t)\end{aligned} \tag{31}$$

Therefore, the mean trajectory is

$$\mathrm{E}[\Delta\mu(n+1)] = \mathrm{E}[\Delta\mu(n)](1 - \gamma R_{\dot{x}\dot{x}}(\Delta\mu(n))), \tag{32}$$

with $\mathrm{E}[\dot{x}n_i] = 0$. Considering that $R_{\dot{x}\dot{x}} = -\ddot{R}_{xx}$, and since we supposed $\Delta\mu(n)$ small, so that $\ddot{R}_{xx}(\Delta\mu(n)) \simeq \ddot{R}_{xx}(0)$, the previous equation can be solved recursively and we get

$$\mathrm{E}[\Delta\mu(n+1)] \simeq \left(1 + \gamma\ddot{R}_{xx}(0)\right)^n \mathrm{E}[\Delta\mu(0)], \tag{33}$$

with $\ddot{R}_{xx}(0) < 0$. This means that $\mu(n)$ converges exponentially to Δ, with a time constant $t_c = -1/\log\left(1 + \gamma\ddot{R}_{xx}(0)\right) \simeq 1/\left(\gamma\ddot{R}_{xx}(0)\right)$. This is illustrated in Fig. 17, where are presented the settling times for the algorithm measured as the rise from 0 to 95% of the final value, $t_s = 3t_c$. This figure shows that the convergence speed is clearly independent of the interpolator and the simulation results are in line with the above development.

We also examined the variation of the convergence speed with respect to the values Δ_1, Δ_2. The results reported in Fig. 18 clearly indicates that the convergence speed is independent of the final values.

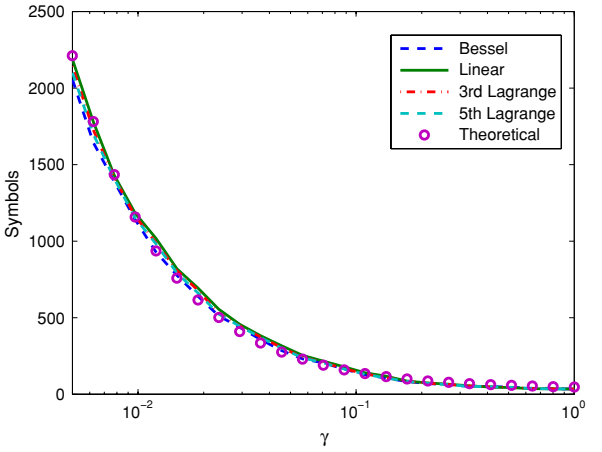


Fig. 17 Settling times for the different interpolators, as a function of the adaptation step, for $\Delta_1 = 0.36T$ and $\Delta_2 = 0.12T$

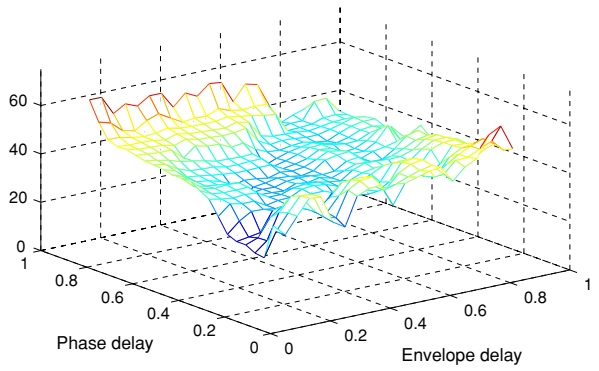


Fig. 18 Settling time for all possible delays (less that T), with $\gamma = 0.2$

5.2 Bias

A careful examination of the results shows that both envelope and phase are often affected by a small, but undeniable, bias. In addition, we have observed in practice that the quadratic error associated with the biased estimates is lower than the error obtained with the true values. Similarly, it appears that the spectrum with identified values is slightly better than the spectrum associated with the true values. This can be explained as follow. The error $e(t)$ is

$$\begin{aligned}e(t) &= x(t) - \tilde{x}(t - \Delta + \mu(n)) \\ &= (x(t) - x(t - \Delta + \mu(n))) + n_i(t),\end{aligned} \tag{34}$$

so that its variance is

$$\begin{aligned}\sigma_e^2 =& \mathrm{E}\left[e(t)^2\right] \\ =& \mathrm{E}[(x(t) - x(t - \Delta + \mu(n)))] + \mathrm{E}\left[n_i(t)^2\right],\end{aligned} \tag{35}$$

that is

$$\sigma_e^2 = \mathrm{E}\left[e(t)^2\right] = 2(R_{xx}(0) - R_{xx}(\mu(n) - \Delta)) + \sigma_i(\mu(n)), \quad (36)$$

where $\sigma_i(\mu(n)) = \mathrm{E}\left[n_i(t)^2\right]$ is the variance of the interpolation noise that depends on the value of the advance $\mu(n)$, see the discussion in § 4.4. The purpose of the algorithm is to find a value μ minimizing σ_e^2. The first term clearly decreases when $\mu(n) \to \Delta$ while the second term may increase. Therefore, the procedure will find the best values that minimize the sum of the two terms, realizing a trade-off between bias and variance of the interpolation noise. The solution is theoretically given by

$$\dot{R}_{xx}(\mu(n) - \Delta) = -\frac{d\sigma_i^2(\mu(n))}{d\mu(n)}. \quad (37)$$

Figure 19 presents the bias measured for the envelope and phase component with respect to the adaptation parameter. We find that the bias still exist and do not depend on the adaptation parameter γ as it was indicated by the analysis. The bias is a function of the interpolator type and order.

5.3 Variance

Variations after convergence are also an important issue. These variations can be modeled as an additive noise on the estimates. With our toy model, we obtained expression (31) for the trajectory of the algorithm. This expression is only valid, on average, if $\Delta\mu(n+1)$ has a zero mean. Thus, taking now $\Delta\mu(n+1) = \mu(n+1) - \bar{\mu}$, where $\bar{\mu}$ is the value at convergence and rearranging the terms, we get

$$\Delta\mu(n+1) = \Delta\mu(n)(1 - \gamma\dot{x}(t + \Delta\mu(n))\dot{x}(t)) + \gamma\dot{x}(t + \Delta\mu(n))n_i(t) = A_n\Delta\mu(n) + B_n n_i(t) \quad (38)$$

Taking the square and the expectation and making the, clearly false, hypothesis that A_n and $\Delta\mu(n)$, B_n and $n_i(t)$ are uncorrelated, we obtain

$$\sigma_\mu^2(n+1) = \mathrm{E}\left[A_n^2\right]\sigma_\mu^2(n) + \mathrm{E}\left[B_n^2\right]\sigma_i^2 \quad (39)$$

with $\sigma_\mu^2(n)$ the variance of μ at step n. We also find $\mathrm{E}\left[A_n^2\right] \simeq 1 + 2\gamma\ddot{R}_{xx}(0) + \gamma^2\mathrm{E}\left[\dot{x}(t)^4\right]$ (with $\mu - \Delta \simeq 0$), and $\mathrm{E}\left[B_n^2\right] = -\gamma^2\ddot{R}_{xx}(0)$. At convergence, we finally obtain

$$\sigma_\mu^2 = \frac{\mathrm{E}\left[\mathrm{B_n^2}\right]\sigma_\mathrm{i}^2}{1 - \mathrm{E}\left[A_n^2\right]} = \frac{\gamma\sigma_i^2}{2 + \gamma\mathrm{E}\left[\dot{x}(t)^4\right]/\ddot{R}_{xx}(0)}. \quad (40)$$

This expression shows that the variance of the estimates increases linearly with γ and that it is only associated to the 'self-noise' in the system. Despite using approximations in the derivation, the simulation results are once more in line with this formula. Figure 20 gives the variance of the estimates of the envelope and phase delays as a function of γ. The linear dependence can be clearly seen.

5.4 EVM—Quadratic error

The EVM is a crucial parameter in digital communications. As we will observe, it sums up some of the other properties. In the case of a single carrier modulation, the EVM is defined by $\mathrm{E}\left[|Z_e - Z_o|^2\right]/\mathrm{E}\left[|Z_o|^2\right]$ where $Z_e = \rho_e e^{j\phi_e}$ and $Z_o = \rho_o e^{j\phi_o}$ stand for the complex envelope of the emitted signal and the original signal respectively. Let us insist on the fact that the situation is more complicated for multi-carrier modulations. In our setting, the expression becomes

$$\mathrm{EVM} = \frac{\mathrm{E}\left[\rho_\mathrm{o}^2 + \rho_\mathrm{e}^2 - 2\rho_\mathrm{o}\rho_\mathrm{e}\cos(\phi_\mathrm{o} - \phi_\mathrm{e})\right]}{\mathrm{E}\left[\rho_o^2\right]}. \quad (41)$$

Therefore, if $\Delta\phi = \phi_o - \phi_e$ is small, $\cos\Delta\phi \simeq 1 - \Delta\phi^2/2$, the expression becomes

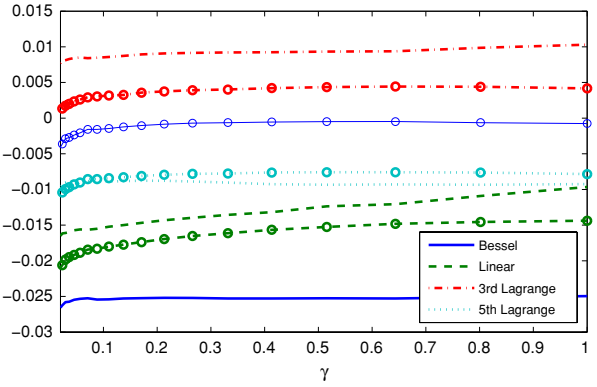


Fig. 19 Bias $\mu - \Delta$ for the different interpolators, as a function of the adaptation parameter γ. Phase (indicated by $-\bigcirc-$) and envelope bias are almost constant with respect to γ

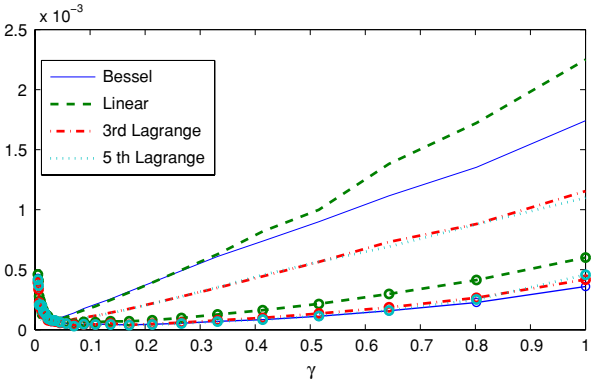


Fig. 20 Variance σ_μ^2 of delays estimates, $\Delta_1 = 0.36T, \Delta_2 = 0.12T$, for the different interpolators, as a function of the adaptation parameter γ. Phase (indicated by $-\bigcirc-$) and envelope variances increase linearly with γ

Analog Integr Circ Sig Process (2008) 55:21–35 33

$$\mathrm{EVM} = \frac{\mathrm{E}\left[(\rho_\mathrm{o} - \rho_\mathrm{e})^2 + \rho_\mathrm{o}\rho_\mathrm{e}\Delta\phi^2\right]}{\mathrm{E}\left[\rho_o^2\right]} \simeq \frac{\mathrm{E}\left[\Delta\rho^2 + \rho_\mathrm{o}^2\Delta\phi^2\right]}{\mathrm{E}\left[\rho_o^2\right]} \tag{42}$$

or

$$\mathrm{EVM} \simeq \frac{\mathrm{E}[\Delta\rho^2]}{\mathrm{E}\left[\rho_o^2\right]} + \mathrm{E}\left[\Delta\phi^2\right] \tag{43}$$

assuming that ρ_o and $\Delta\phi$ are independent. As a result, the EVM is a function of the quantities we previously studied. We already established that the estimates are biased and that their variances, which are proportional to interpolation noise, increase linearly with γ. Therefore, the EVM behaves as

$$\mathrm{EVM} \simeq \frac{\gamma}{2}\left[\frac{\sigma_\rho^2}{\mathrm{E}\left[\rho_o^2\right]} + \sigma_\phi^2\right] + \left[\frac{B_\rho^2}{\mathrm{E}\left[\rho_o^2\right]} + B_\phi^2\right], \tag{44}$$

where σ_ρ^2 and σ_ϕ^2 are the variances of interpolation noise for the envelope and phase, and B_ρ and B_ϕ the bias terms. Since $\sqrt{1+\alpha x} \simeq 1 + \alpha x/2$, the EVM rms may also show the same behaviour. This is the case in Fig. 21 where we clearly see a linear slope and an offset value. It also shows that the Lagrange and Bessel interpolators have the best performances.

It is also interesting to look at the EVM with respect to all possible delays and with a fixed γ. Figure 22 presents these results for $\gamma = 0.05$. We can observe a 'modulation' which is related to the variable variance of the interpolation noise with respect to the delay (see Fig. 14). Here, as in previous examples, we took $T/Ts = 5$ and indeed notice error minima at interpolation nodes.

5.5 Summary of the results

In the previous developments we have examined the behaviour of our algorithm according to the choice of interpolators and adaptation steps. Using specific examples, we have demonstrated a vast improvement can be achieved in terms of EVM and output spectrum, as shown in Figs. 15 and 16. Furthermore, using an approximate 'toy' model of the algorithm, we have shown theoretically and checked numerically, that

- the settling time decreases exponentially with the adaptation step (Fig. 17),
- a small estimation bias exists that does not depend on the adaptation parameter but rather of interpolation order (Fig. 19). The presence of this bias is explained as a way to reduce the whole quadratic error,
- the variance (40) behaves approximately as a linear function of the adaptation parameter (Fig. 20),
- the EVM (44) increases linearly with the adaptation step γ (Fig. 21).

As a consequence, a trade-off has to be made between the two important parameters that are the settling time and the EVM. The choice of the order of the interpolation is also important with respect to performances and implementation costs. The main values for the bias and variance can be found in Table 2 for different interpolators and several values of the adaptation step γ. The results for the settling time and the EVM are reported in Table 3 for the same conditions. In order to ensure independence in signal scales, the signal power was normalized to one.

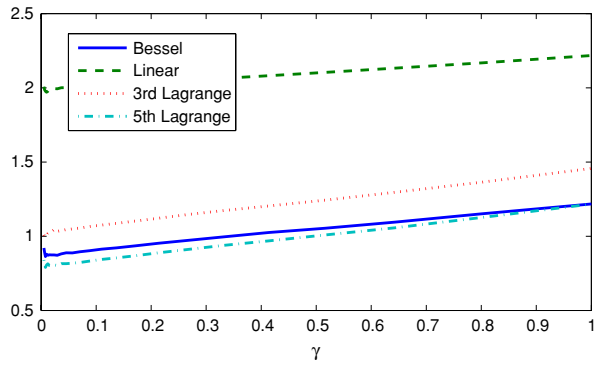


Fig. 21 EVM rms (in %) at the output of the algorithm, for delays $\Delta_1 = 0.36T, \Delta_2 = 0.12T$ as a function of the adaptation parameter γ

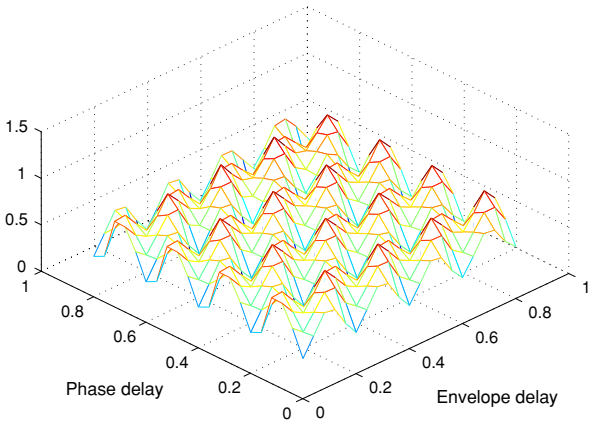


Fig. 22 EVM in percents at output of the algorithm with $\gamma = 0.05$, for all possible delays. The 'modulation' is related to the non-stationary interpolation noise that vanishes at interpolation nodes

6 Conclusion

In this article, we proposed and analysed a new synchronization algorithm tailored for an EER architecture. Performances are indeed severely degraded in case of delay mismatch between envelope and phase paths. The proposed

Table 2 Bias ($\mu - \Delta$) and variance σ_μ^2 of the estimates of delays, for different interpolators and several values of the adaptation step γ

γ		Bias $\times 10^3$			Variance $\times 10^5$		
		3rd Lag	Bessel	5th Lag	3rd Lag	Bessel	5th Lag
0.05	ρ	−15.5	8	−9	6.2	10	5.8
0.05	ϕ	−19	2	−9	5.8	6.59	5.75
0.1	ρ	−15	8	−8	11.5	21	11.5
0.1	ϕ	−28	4	−8	4.5	4.6	4.5
0.5	ρ	−14	8	−7	56.5	89	56
0.5	ϕ	−17	4	−9	13	11.5	11.5

Table 3 Settling times of the algorithm and EVM rms (in %) for the output, for different interpolators and several values of the adaptation step γ

γ	Settling time	EVM rms %		
		3rd Lag	Bessel	5th Lag
0.05	228	1.045	0.89	0.815
0.1	127	1.07	0.9	0.84
0.5	45	1.24	1.05	1.009

algorithm is based on an adaptive LMS structure. We studied the influence of the adaptation on the quadratic error, EVM and output spectrum. The algorithm uses an interpolation procedure that is characterized in terms of performances and in terms of implementation. We also examined the influence of the sampling rate. These studies demonstrated the strong interest of such a mandatory correction algorithm and the enhancement of transmitter performances.

Regarding the algorithm, a few points can be further investigated: adopting different adaptation steps for the two recursions, decreasing the number of adaptation steps and alternating the two recursions. Higher system view point has to be taken into account in order to define the best use of the procedure: full tracking of delays or training periods. A for the implementation, modified and optimized Farrow structures should be considered.

This solution has drawbacks. For instance, it requires I/Q demodulation and uses a Cordic processor to split the envelope and phase of the PA output signal (derivative computation). We are currently working on these issues to improve our system.

Other concerns are the behaviour of the system with further mismatches such as a complex gain in the feedback loop, non-linear distortions and quantization noise. Preliminary results show that the algorithm is robust against the gain and noise mismatch, in the sense that the optimum solution remains unchanged. Evidently further mismatches imply a degradation of performances so future work will include attempts to also correct these mismatches with the same feedback loop.

Acknowledgements The authors are indebted to the anonymous referees for their careful reading of the manuscript as well as their valuable comments and suggestions that helped improve the presentation of this article. Thanks are extended to S. Trudel, P. J. Duriez and L. Garbell for their invaluable and friendly proofreading of the manuscript.

References

1. Abramowitz, M., & Stegun, I. A. (1965). *Handbook of mathematical functions: With formulas, graphs, and mathematical tables*. Dover Publications.
2. Baudoin, G., Berland, C., Villegas, M., & Diet, A. (2003). Influence of time and processing mismatches between phase and envelope signals in linearization systems using envelope elimination and restoration, application to hiperlan2. In *Microwave Symposium Digest, 2003 IEEE MTT-S International* (Vol. 3, pp. 2149–2152.
3. Bercher, J. F, & Berland, C. (2006). Envelope/phase delays correction in an EER radio architecture. In *Proceedings of the 13th IEEE International Conference on Electronics, Circuits, and Systems, ICECS2006* (pp. 443–446).
4. Bercher, J. F., Diet, A., Berland, C., & Baudoin, G., & Villegas, M. (2004). Monte-carlo estimation of time mismatch effect in an OFDM EER architecture. In *Proceedings of the 2004 IEEE Radio and Wireless Conference* (pp. 283–286).
5. Candan, C. (2007) An efficient filtering structure for Lagrange interpolation. *IEEE Signal Processing Letters, 14*, 17–19.
6. Crochiere, R., & Rabiner, L. (1975). Optimum fir digital filter implementations for decimation, interpolation, and narrow-band filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 23*(5), 444–456.
7. Dempster, A., & Murphy, N. (2000). Efficient interpolators and filter banks using multiplier blocks. *IEEE Transactions on Signal Processing, 48*, 257–261.
8. Farrow, C. W. (1988). A continuous variable digital delay element. In *Proceedings of IEEE International Symposium Circuits Systems* (pp. 2641–2645).
9. Hentschel, T. (2002). *Sample rate conversion in software configurable radios*. Artech House Publishers.
10. Hentschel, T., & Fettweis, G. (2002). *The digital front-end –bridge between rf-and baseband-processing*. Software Defined Radio Enabling Technologies, John Wiley & Sons Ltd, pp. 151–198.
11. Jau, J. K., & Horng, T. S. (2001). Linear interpolation scheme for compensation of path delay difference in an envelope elimination and restoration transmitter. In *Proceedings of Asia-Pacific Microwave Conference, 2001. APMC 2001* (Vol. 3, pp. 1072–1075).
12. Johansson, H., & Gustafsson, O. (2005). Linear-phase FIR interpolation, decimation, and *m*th-band filters utilizing the Farrow structure. *IEEE Transactions on Circuits and Systems I, 52*, 2197–2207.
13. Johansson, H., & Lowenborg, P. (2003). On the design of adjustable fractional delay FIR filters. *IEEE Transactions on Circuits and Systems II, 50*, 164–169.
14. Kahn, L. (1952). Single-sideband transmission by envelope elimination and restoration. *Proceedings of the IRE, 40*(7), 803–806.
15. Laakso, T. I., Valimaki, V., Karjalainen, M., & Laine, U. K. (1996). Splitting the unit delay. *IEEE Signal Processing Magazine, 13*, 30–60.

16. Makundi, M., Laakso, T., & Valimaki, V. (2001). Efficient tunable IIR and allpass filter structures. *Electronics Letters, 37*, 344–345.
17. Mártires, J., Borg, C., & Larsen, T. (2006) Differential delay equalization in a Kahn EER transmitter. In *ST Mobile & Wireless Communications Summit*, nr. 15th, Myconos, Grækenland.
18. Morgan, P. (2006). Highly integrated transceiver enables high-volume production of GSM/EDGE handsets. *RFdesign* (pp. 36–42).
19. Oetken, G. (1979). A new approach for the design of digital interpolating filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 27*(6), 637–643.
20. Raab, F. (1977). Idealized operation of the class E tuned power amplifier. *Circuits and Systems, IEEE Transactions on Circuits and Systems, 24*, 725–735.
21. Sander, W. B., Schell, S.V., & Sander, B. L. (2003). Polar modulator for multi-mode cell phones. *Proceedings of the IEEE 2003 Custom Integrated Circuits Conference* (pp. 439–445).
22. Sarigeorgidis, K., & Rabaey, J. (2004). Ultra low power cordic processor for wireless communication algorithms. *Journal of VLSI Signal processing, 38*, 115–130.
23. Staszewski, R., Wallberg, J., Rezeq, S., Hung, C. M., Eliezer, O., Vemulapalli, S., Fernando, C., Maggio, K., Staszewski, R., Barton, N., Lee, M. C., Cruise, P., Entezari, M., Muhammad, K., & Leipold, D. (2005). All-digital PLL and transmitter for mobile phones. *IEEE Journal of Solid-State Circuits, 40*(12), 2469–2482.
24. Strasser, G., Lindner, B., Maurer, L., Hueber, G., & Springer, A. (2006). On the spectral regrowth in polar transmitters. In *Microwave Symposium Digest, 2006. IEEE MTT-S International* (pp. 781–784).
25. Valimaki, V. (1995) A new filter implementation strategy for Lagrange interpolation. In *Proceedings of the 1995 IEEE International Symposium on Circuits and Systems, ISCAS '95* (Vol. 1, pp. 361–364).
26. Valimaki, V., & Laakso, T. I. (2000). Principles of fractional delay filters. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'00 6*.
27. Vesma, J., & Saramaki, T. (1997). Optimization and efficient implementation of FIR filters with adjustable fractional delay. In *Proceedings of 1997 IEEE International Symposium on Circuits and Systems, ISCAS '97* (Vol. 4, pp. 2256–2259).
28. Volkov, E. A. (1986). *Numerical methods*. Mir Publisher Moscow.
29. Whittaker, J. M. (1935). *Interpolatory function theory*. Cambridge, England: Cambridge University Press.
30. Yli-Kaakinen, S. (2006). Multiplication-free polynomial-based FIR filters with an adjustable fractional delay. *Circuits, Systems, and Signal Processing, 25*, 265–294.
31. Yli-Kaakinen, J., & Saramaki, T. (2004). An algorithm for the optimization of adjustable fractional-delay all-pass filters. In *Proceedings of the 2004 International Symposium on Circuits and Systems ISCAS'04* (Vol. 3, pp. III-153–156).
32. Yli-Kaakinen, J., & Saramaki, T. (2007). A simplified structure for FIR filters with an adjustable fractional delay. *IEEE International Symposium on Circuits and Systems, 2007 ISCAS 2007* (pp. 3439–3442).

Jean-François Bercher was born in Marseille, France, in 1965. He is maître es Sciences et Techniques in biomedical engineering (1988) and Engineer of the Institut National Polytechnique de Grenoble (1990). He received the Doctorat degree in physics at the Laboratoire des Signaux et Systèmes, Université de Paris-Sud, Orsay, in 1995. He is now with the University of Paris-Est, ESIEE-Paris, where he is Associate Professor. His main interests are statistical radiotechnique and applications of information theory to signal processing and telecommunications.

Corinne Berland Corinne Berland was born in France, in 1966. She is engineer of the ESIEE (1989) and received the Doctorat degree in physics at the University of Marne La Valle in 2001. She worked until 1998 in Alcatel where she was in charge of RF asics development for GSM handportable. She is now with the University of Paris-Est, Electronics Systems Department of ESIEE and ESYCOM ESIEE, where she is Associate Professor. Her main interests are transceiver architecture for radiocommunication systems and RF/analog integrated circuits.

8.13 J.-F. Bercher and C. Berland, "Adaptive delays alignment in polar transmitter architecture," submitted to *IEEE trans. on Circuits and Systems*, april 2009.

Adaptive delays alignment in polar transmitter architecture

J.-F. Bercher[1] and C. Berland[2]

[1] Université Paris-Est, LabInfo-IGM, ESIEE-Paris
[2] Université Paris-Est, Département Systèmes électroniques, ESIEE Paris
Cité Descartes, BP99, 93162 Noisy le Grand cedex
{jf.bercher,c.berland}@esiee.fr

***Abstract*—Polar architecture is a regarded solution for the design of efficient transmitters for radiocommunication systems. However, this architecture is sensitive to delay mismatches between the envelope and the phase paths. In this paper, we propose to address this problem by baseband digital signal processing. We present several adaptive algorithms based on the minimization of the mean square error between the ideal signal and the signal generated by the transmitter. Since the system gain and phase offset of the system are not precisely known, we also consider the problem of the identification of these quantities in the feedback loop. We propose solutions for the direct correction of the unknown delays, which also account for the gain variation and phase offset. A model of a polar architecture is tuned on Agilent Advanced Design System and used to evaluate the algorithms. Results assess both the need of a calibration procedure and the relevance of the proposed solutions.**

***Index Terms*—Polar transmitters, EER architecture, time alignment, synchronization, baseband signal processing, least mean square methods**

EDICS:
COMM200 - Wireless Communication Systems
SIPRO100 - Adaptive signal processing
COMM260 - Circuits and Systems for Software-Defined Radio

I. Introduction

Since new applications and upcoming standards require increasing data rates and lower power consumption, the study and development of new architectures of communication transmitters for user units is very active. In particular, architectures relying on the polar representation of signals, called polar transmitters, are highly regarded. Indeed, the polar architecture is a solution to achieve a very efficient and linear transmitter [1]. Further, it is flexible enough to enable the development of multiband [2] and multimode solutions [3], [4] for Software-Defined Radio (SDR). This architecture is based on Kahn's Envelope Elimination and Restoration (EER) procedure [5]. It has been improved so that actual solution tends toward a fully digital transmitter [6]. The phase-modulated signal is translated to RF frequency using either a classical I/Q modulator or through a modulated phase locked loop. The envelope is restored at the output of the transmitter by varying the supply of the power amplifier (PA) through a switching power supply.

The drawback of this architecture is its sensitivity to the difference of propagation and processing delays between amplitude and phase signals paths. The mismatch between the two paths entails spectral regrowths [7], [8] and an increase of the Error Vector Magnitude (EVM) [9]. Fulfillment of the strict requirements on transceivers for modern modulation schemes warrants a high precision timing alignment between the envelope and phase path [10]. This becomes more and more critical and challenging as the data rate increases [10], [11], [12].

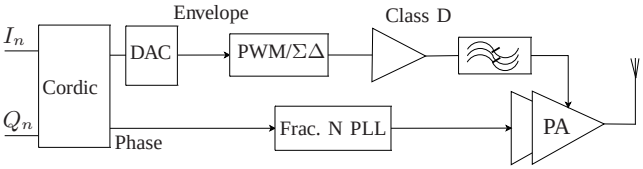


Fig. 1. Polar transmitter. Instead of the classical I-Q decomposition, the polar transmitter rely on the polar decomposition of signals. Here a CORDIC processor (Coordinate Rotation Digital Computer) is used. The phase-modulated signal is translated to RF frequency using a modulated phase locked loop, either digital or analog. The envelope is restored at the output of the transmitter while varying the supply of the power amplifier (PA) through a switching power supply. A drawback of the architecture is its sensitivity to the differential processing delay between the two paths.

This mismatch is due to pipeline differences in the paths and to the delay in the anti-alias filter (amplitude path), as well as small contributions from other analog delays [13]. The asymmetry of processing, the independence of underlying clocks and the analog interfaces also imply delay mismatches in the recent fully digital solutions, which also require a precise delay alignment of amplitude and phase paths [14]. In a production environment, even with the global knowledge of various delays in the transmitter path, these delays have to be tuned on a case-by-case basis to include process spreading on elements. Overall, this mismatch becomes more and more limiting as data rate increases. Furthermore, current devices experience severe variations of temperature, modifications of the supply voltage, output frequency and power [15] that impact the system behavior and can also imply further time-varying desynchronizations that, according to their amount, would need a tracking procedure.

A linear interpolation was suggested in [16] so as to compensate for the mismatch, while a group delay equalizer was proposed in [17]. In these two cases, it still remains to identify the delay mismatch and track its possible variations. An adaptive hardware realization was proposed in [18], but

with a precision limited by the underlying clock. A calibration procedure, based on a feedback loop and the maximization of the input-output covariance function, is described in [19].
We elaborate here on the principle of an adaptive delay correction presented in [20], [21]. We propose solutions that cover the correction and tracking of the unknown delays, with high accuracy and low complexity. We also show that the solutions are robust to further mismatches, such as uncertainty in the knowledge of gain and phase shift of the transmitter. In order to lower the residual noise and improve the performances, we propose a solution for the identification of the complex gain. The effectiveness of the solution is illustrated on the case of the 3GPP standard for user equipment (UE) transmitters.
The paper is organized as follows. In section II, we present the model of the architecture considered, we illustrate the need of envelope and phase alignment, and discuss its feasibility in polar based transmitter architectures. We also analyze the main characteristics of the mean square error between the input and output of the system. In section III, we consider the problem of the delays correction, based on the minimization of the input-output means square error. We propose and argue a correction algorithm, implemented using a stochastic gradient. In order to reduce its complexity, we also describe a suboptimal version. The performances of these two solutions are illustrated through simulations. In section IV, we consider the behavior of these solutions subject to uncertainties in the knowledge of system gain and phase offset. In order to preserve good performances, we introduce an identification of these unknowns in the feedback loop. The efficiency of the whole solution, in terms of convergence speed, error vector magnitude and spectral regrowths, is then demonstrated through simulation results. We finally conclude in section V.

II. Envelope and phase alignment

In this section, we first indicate that envelope and phase alignment is an essential issue in the design of polar transmitters for modern systems. We discuss the feasibility of the alignment of the envelope and phase components.Then, we examine the behavior of the mean square error between the original signal and the signal at the output of the transmitter, and consider the minimization of this mean square error for the compensation of the delay mismatches.

A. Simulation model of the transmitter – Impact of delay mismatches on a 16QAM modulation

In this paper, we will illustrate the issues and the performances of proposed algorithms on a 16QAM modulation with a data rate of 3.84 Mcps. This corresponds to the 3GPP standard [22]. Transmitter characteristics are given in term of EVM (Error Vector Magnitude), spectrum transmitter mask and ACLR. We will mainly focus here on EVM and output spectrum. The output spectrum is specified relative to the root-raised cosine filtered mean power of the UE carrier taken at different frequency offsets from the carrier. Specifications are as follows:

- at 2.5 MHz, -35 dBc in 30 kHz bandwidth,
- at 3.5 MHz, -50 dBc in 30 kHz bandwidth,
- at 4 MHz, -35.5 dBc in 1 MHz bandwidth,
- for F>8 MHz, -49 dBc in 1 MHz bandwidth.

The EVM has to remain lower than 17.5%.
The transmitter model, see Fig. 2, is elaborated with Agilent Advanced Design System (ADS). The complex modulated signal is generated with a 16QAM modulator, followed by a 0.22 roll-off root-raised cosine filter. A complex to polar conversion is then realized, with a quantization on 10 bits at a rate $T_s/12$, where T_s denotes the symbol period. Conversions to the analog domain are realized using oversampling DACs followed by baseband filters. The processing delays, as well as delays introduced by filters and other analog elements, are modeled as two delays Δ_1 and Δ_2 in the envelope and phase paths respectively. As we will describe in the following sections, the correction algorithms involve the comparison between the ideal signal and the transmitted one. Hence, a return path has to be implemented, using either an I/Q demodulator or a simple frequency translation (depending on the algorithm) and an analog to digital conversion. The signal is quantified on 10 bits and the converter uses the same clock as in the transmit path, that is to say with a sampling rate of $T_s/12$. Finally, we introduce a complex gain to account for the gain and phase rotation of the modulated signal after the power amplifier. In this work we do not elaborate further on the model of the power amplifier and only take into account the gain and phase rotation. Potential nonlinear effects, especially, are not considered here where we concentrate on the issues related to time mismatches. The calibration and correction of nonlinearities, which usually involves a feedback loop, is well documented, but still an open issue [23].
With this model, without gain, phase rotation or time mismatches, the output EVM is on the order of 0.3%, due to the quantification on 10 bits. Spectral regrowth at 3.5 MHz from the carrier is at -70 dBc.
In order to examine the sensitivity of polar architecture to delay mismatches between phase and envelope paths, we introduced into the simulation model delay mismatches from $T_s/2$ to $T_s/50$. Impacts of these mismatches on the EVM value and the output spectrum are reported Table I.
According to specifications such as 3G specifications on the EVM and spectrum, a maximum acceptable delay shall be on the order of $T_s/50$, with the symbol period $T_s = 0.26\ \mu s$, that is to say lower than 5 ns. These results demonstrate that a synchronization algorithm is mandatory.

Delay	**EVM (%)**	**Relative Spectrum @**$3.5\ MHz$
$T_s/2$	77%	-33 dBc
$T_s/10$	12.4%	-48 dBc
$T_s/20$	6.5%	-55 dBc
$T_s/50$	3.6%	-61 dBc

TABLE I
Impacts of delay mismatches on the EVM and on the spectrum.

B. Feasibility of the time alignment

Let $X(t)$ be the complex envelope of the signal at the output of the digital modulator

$$X(t) = \rho(t)\exp(j\phi(t)) = \rho(t)\cos(\phi(t)) + j\rho(t)\sin(\phi(t)),$$

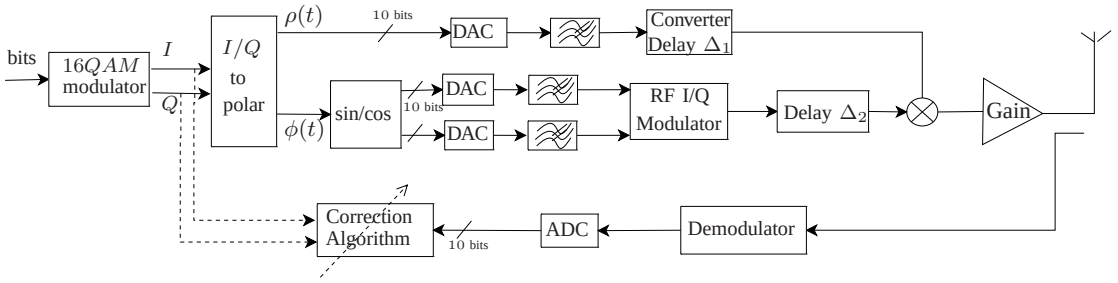


Fig. 2. Simulation model of the transmitter. This model includes an I/Q modulator followed by an I/Q to polar converter. The envelope signal is converted into analog using a 10 bits DAC. In the envelope path, the impact of the DC-DC converter is modeled as a delay Δ_1. The phase signal is translated to RF with an RF I/Q modulator after conversions by 10 bits DACs. The processing delay is introduced as a delay Δ_2 before the restoration of the complete modulated signal. A complex gain in the return path accounts for gain and phase shifts that affect the emitted signal.

where $\rho(t)$ and $\phi(t)$ denote the envelope and phase signals, as shown in Fig. 2.

This baseband signal $X(t)$ is processed by the RF transmitter (shaping, RF translation, amplification) which results in the signal at the antenna, whose baseband complex envelope is denoted $Z(t)$. This signal is affected by the delays introduced by the RF processing, and the envelope and phase components are delayed respectively by Δ_1 and Δ_2. The signal $Z(t)$ then becomes

$$Z(t) = \rho(t-\Delta_1)\exp\left(j\phi(t-\Delta_2)\right).$$

For complex modulation schemes, $X(t)$ can often be modelled as a complex Gaussian process. In the case of a complex circular Gaussian process it is well known that envelope $\rho(t)$ and phase $\phi(t)$ are independent and respectively distributed according to Rayleigh and uniform distributions.

The case of delayed envelope and phase is less known. In fact, it appears that $\rho(t-\Delta_1)$ and $\phi(t-\Delta_2)$ are also Rayleigh and uniform independent variables, with no reference to the correlation structure of the underlying original Gaussian signal $X(t)$, for any delays Δ_1, Δ_2, see [9]. This result indicates that for Gaussian processes, envelope and phase are *always* independent whatever the delay between envelope and phase components. This shows that the statistics of the output at a single instant do not convey any information on the time alignment or mismatch between envelope and phase components. Hence, any solution must involve statistical quantities computed from several instants. A simple example is the correlation function, or equivalently the power spectrum. If $z(t) = \mathrm{Re}\{Z(t)\}$ denotes the real part of the output, its correlation function is

$$\begin{aligned} R_{zz}(\tau) = \mathrm{E}\{&\rho(t-\Delta_1)\cos(\phi(t-\Delta_2)) \\ &\times\rho(t-\tau-\Delta_1)\cos(\phi(t-\tau-\Delta_2))\} \end{aligned} \quad (1)$$

where $\mathrm{E}[\bullet]$ denotes the statistical expectation operator. The point that we must emphasize here is that we do not have a closed-form expression for $R_{zz}(\tau)$, unfortunately, and it even does not seem possible to obtain a workable series representation. Series expansion exist for the correlation of the $\cos(\phi)$ term, or for the correlation of the envelope [24]. Consequently it is difficult to quantify the dependence of the correlation function on the delay mismatch and imagine a procedure relying only on this correlation function. However, if we introduce a delay μ in the envelope path (for instance), denote $\Delta = \Delta_2 - \Delta_1$ and use the stationarity of the signal, we obtain

$$\begin{aligned} R_{zz}(\tau,\mu) = \mathrm{E}\{&\rho(t-\mu)\cos(\phi(t-\Delta)) \\ &\times\rho(t-\mu-\tau)\cos(\phi(t-\Delta-\tau))\}. \end{aligned} \quad (2)$$

Then we can observe that this new correlation function is exactly equal to the correlation function of the original aligned signal $R_{xx}(\tau)$ when $\mu = \Delta$. Therefore, a possible procedure could be to look (numerically) fora value of μ that minimizes some distance between $R_{zz}(\tau,\mu)$ and $R_{xx}(\tau)$, for any τ, e.g.

$$\int \left(R_{zz}(\tau,\mu) - R_{xx}(\tau)\right)^2 \mathrm{d}\tau, \quad (3)$$

or equivalently in the frequency domain:

$$\int \left(S_{zz}(f,\mu) - S_{xx}(\tau)\right)^2 \mathrm{d}f. \quad (4)$$

From these two procedures, we can yet observe that (i) we need a feedback from the output, that is a way of acquiring and computing some quantity related to the output of the system, and (ii) we need to compare it to some known or estimated characteristic of the input signal. This means that we shall implement a feedback loop, relying on the demodulation and digitization of an image of the emitted signal taken with a coupler. It shall be observed that a feedback loop is often already present for calibration or tracking purposes (namely for the power amplifier calibration and control).

But the suggestions above are actually difficult to implement because they require the evaluation of (integrals or sums of) correlation functions, or spectra, which is computationally expensive and statistically delicate; the control algorithm can not be derived precisely, nor characterized, because of the lack of closed-form expression of the correlation.

Since we see that a feedback loop is almost unavoidable, our proposal is to consider the direct error between the input and output signals and examine the possibility to derive compensation procedures of the delay mismatches based on this error.

C. The mean square error between the original and generated signals

As above-mentioned, we note $X(t)$ the complex envelope of the signal at the output of the digital modulator and $Z(t)$ the

baseband complex envelope of the signal at the antenna. We define by

$$J(\Delta_1, \Delta_2) = \mathrm{E}\left[|e(t)|^2\right] = \mathrm{E}\left[|x(t) - z(t)|^2\right] \quad (5)$$

the mean square error between these two signals, with $x(t) = Re\{X(t)\}$ and $z(t) = Re\{Z(t)\}$.
Taking into account the independence between $\rho(t_1)$ and $\phi(t_2)$, for any instants t_1, t_2, it reduces to

$$J(\Delta_1, \Delta_2) = 4R(0,0) - 4R(\Delta_1, \Delta_2) \quad (6)$$

where

$$R(\Delta_1, \Delta_2) = \mathrm{E}\left[\rho(t)\cos(\phi(t))\rho(t-\Delta_1)\cos(\phi(t-\Delta_2))\right] \quad (7)$$

is a kind of 'correlation function'. As already indicated, we do not have a closed-form expression for $R(\Delta_1, \Delta_2)$. Therefore, we must rely on numerical simulations or approximations to quantify the behavior of the algorithms presented in the following sections.
We can still note here that $R(0,0)$ reduces to $R(0,0) = \mathrm{E}\left[\rho(t)^2\right]\mathrm{E}\left[\cos(\phi(t))^2\right]$, and that $R(\tau,\tau)$ obtained with $\tau = \Delta_1 = \Delta_2$ is nothing else but the correlation function $R_{xx}(\tau)$. Since we know that the correlation function is a function of the shaping filter, it follows that the behavior of the error $J(\Delta_1, \Delta_2)$ is closely related to the shaping filter. We shall also note that $J(\Delta_1, \Delta_2) = \int S_{ee}(f)\mathrm{d}f$, where $S_{ee}(f)$ is the power spectrum of the error $e(t)$. Hence, the mean square error $J(\Delta_1, \Delta_2)$ also represents the power spectrum of the error induced by the mismatches Δ_1 and Δ_2, including the spectral regrowths.
The error function $J(\Delta_1, \Delta_2)$ was evaluated numerically in the case of a 16QAM modulation with a square-root Nyquist filter (root-raised cosine with 0.22 roll-off). It is presented in Fig. 3 for delays lower than 3 symbol periods. We clearly observe a global minimum, but some local minima also appear for more important delays. Therefore, one can define a basin of attraction $\mathcal{C}$, as reported in Figs. 3 and 4 such that any descent algorithm will converge efficiently to the global minimum. This indicates that it is possible to adjust efficiently some parameters on the input or on the output in order to minimize the resulting error. With such an approach, we might be able to account for delays that belong to the domain of convergence in Fig. 4. If this domain of convergence is too restrictive, it is still possible to enlarge it by designing specific training sequences with a smoother error function (but this supposes a pre-calibration approach).
Because of the degradation of performances as discussed before, it is mandatory to counterbalance the relative delay $\Delta_2 - \Delta_1$ between the two components at the antenna. A natural approach is to identify these delays, see [20], and then correct them. But it is also possible to directly compensate the delays without requiring a preliminary identification step.

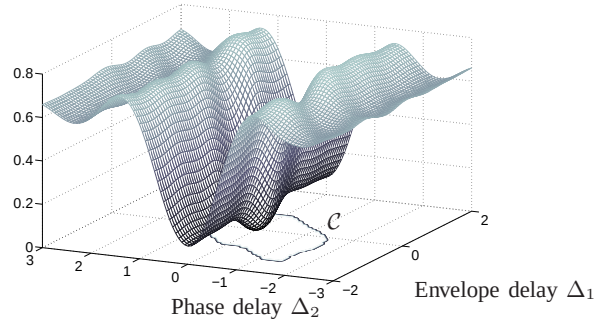


Fig. 3. Shape of criterion for a 16QAM, with the indication of the basin of attraction, such that the minimum can be obtained using a descent algorithm from any initial point in the basin.

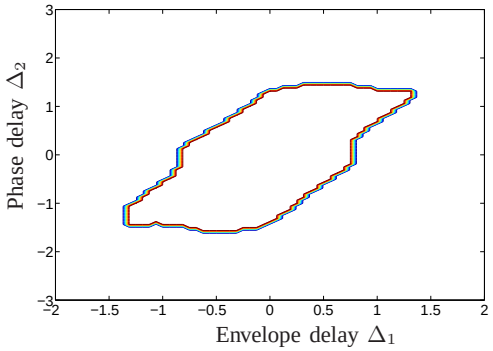


Fig. 4. Domain of convergence $\mathcal{C}$ for a 16QAM with a Nyquist square root shaping filter with roll-off of 0.22, which corresponds to the 3GPP standard. Larger domains are obtained when the roll-off increases (smoother error function).

III. A CORRECTION ALGORITHM

A. Principle and derivation of the correction algorithm

The basic idea is to introduce two advances, say μ_1 and μ_2, as indicated in Fig. 5 in the envelope and phase paths. Then the advances are tuned so as to minimize a distance, e.g. the quadratic distance, between the resulting observation and the original aligned signal: the optimum values of the two advances will exactly compensate the delays introduced by the analog paths.

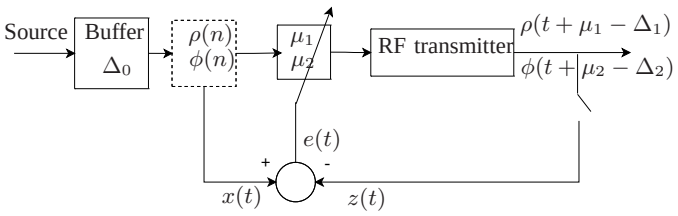


Fig. 5. Principle of the feedback loop for delays correction: two advances μ_1 and μ_2 are introduced in the main path (the original signal is buffered so as to be able to compute these advances). The remaining delays at output are $\Delta_1 - \mu_1$ and $\Delta_2 - \mu_2$. The output is compared to the original aligned input, and μ_1, μ_2 are adjusted so as to minimize this error. The two signals are synchronized when $\mu_1 = \Delta_1$ and $\mu_2 = \Delta_2$.

With the advances μ_1 and μ_2 introduced in the envelope and phase paths, the output becomes

$$z(t) = \rho(t + \mu_1 - \Delta_1)\cos(\phi(t + \mu_2 - \Delta_2)). \quad (8)$$

The related mean square error is

$$J(\mu_1,\mu_2) = \mathrm{E}\big[|e(t)|^2\big] = \mathrm{E}\big[|x(t)-\rho(t_1)\cos(\phi(t_2))|^2\big], \\ = 4R(0,0) - 4R(\Delta_1-\mu_1, \Delta_2-\mu_2). \quad (9)$$

where we noted $t = nT$, with T the sampling period, and also used $t_1 = t+\mu_1-\Delta_1$ and $t_2 = t+\mu_2-\Delta_2$ in order to simplify the expressions.

As already mentioned we do not have a closed-form for this criterion nor, of course, a direct explicit solution for its minimizers. But we can reach the solution with a descent algorithm. For instance we can use a gradient algorithm that consists in iterating

$$\begin{aligned} \mu_1(k+1) &= \mu_1(k) - \gamma_1(k)\left.\frac{\partial J}{\partial \mu_1}\right|_{\mu_1=\mu_1(k)} \\ \mu_2(k+1) &= \mu_2(k) - \gamma_2(k)\left.\frac{\partial J}{\partial \mu_2}\right|_{\mu_2=\mu_2(k)} \end{aligned} \quad (10)$$

where γ_1 and γ_2 are adaptation steps. The gradients are computed according to

$$\frac{\partial J(\mu_1,\mu_2)}{\partial \bullet} = \frac{\partial \mathrm{E}\left[|e(t)|^2\right]}{\partial \bullet} = 2\mathrm{E}\left[\frac{\partial e(t)}{\partial \bullet} e(t)\right]. \quad (11)$$

In practice, we have to resort to an approximation of the theoretical algorithm, using an appropriate approximation of the unknown theoretical expectations. A classical solution is to adopt a stochastic gradient algorithm which consists in using the instantaneous gradient rather than the statistical average, and in updating the equations at each new sample, at the rate T. This leads to the two updating rules for μ_1 and μ_2:

$$\begin{aligned} \mu_1(n+1) &= \mu_1(n) + \gamma_1(n)\left.\frac{d\rho(u)}{du}\right|_{t_1}\cos\phi(t_2)\, e(t) \\ \mu_2(n+1) &= \mu_2(n) + \gamma_2(n)\rho(t_1)\left.\frac{d\cos\phi(u)}{du}\right|_{t_2} e(t) \end{aligned} \quad (12)$$

In this approach, we have to adjust the two advances μ_1 and μ_2 and apply them to the input signal. Of course, working with digital signals implies that the data are only available at sampling points. Adopting very high sampling frequencies in order to get some required precision, is not an efficient solution with respect to cost, power consumption, and realization. A classical solution for tuning delays/advances from available data samples is digital interpolation. In order to be able to compute the advances by digital interpolation, we need to dispose of some samples in the future of the current point. Therefore the procedure includes a buffering of a few samples and the introduction of a small delay, say Δ_0, with respect to the message delivered to the transmitter.

Let us now turn to the description of the interpolation procedure. The task of approximating a function given a series of samples is solved, in numerical analysis, by Lagrange polynomial interpolation. A benefit of polynomial interpolators is that they can be implemented very efficiently in hardware, and that coefficients can be computed in real time. An efficient structure was devised by Farrow [25]. A recent improvement is given in [26]. Furthermore, in such structures, the delay is directly adjustable without modification so that they are very suitable for adaptive synchronization problems. It is worth mentioning that in the last decade, many solutions have been proposed for the synthesis and optimization of adjustable fractional delay filters, especially suitable in the case of large bands or when strong anti-aliasing is needed. In our context, the over sampling ratio is high and the frequency response of the interpolators is very flat in magnitude and linear in phase in the region of interest. Therefore, we simply adopt here a standard Lagrange interpolator, implemented using a Farrow structure. Performances and comparisons with respect to the choice of interpolator and with respect to the order are given in [21]. We select here a 5^{th} order Lagrange interpolator as an interesting trade off between performances and complexity. For this interpolator, if x is a sampled sequence at the rate T, then the interpolated value $x(m,\tau)$ at the time $(m+\tau)T$ is computed by

$$\begin{aligned} x(m,\tau) &= \tfrac{(\tau^2-1)(\tau-2)\tau}{24}x(m-2) \\ &- \tfrac{(\tau^2-4)(\tau-1)\tau}{6}x(m-1) + \tfrac{(\tau^2-1)(\tau^2-4)}{4}x(m) \\ &- \tfrac{(\tau^2-4)(\tau+1)\tau}{6}x(m+1) + \tfrac{(\tau^2-1)(\tau+2)\tau}{24}x(m+2). \end{aligned} \quad (13)$$

The computational load of algorithm (12) is of about 6 real multiplications per iteration. We have to compute the error and two derivatives that can be simply approximated by finite differences. Of course, we also have the cost of interpolation in the two branches, which has a complexity of p^2 for the Farrow structure and of $3p-1$ for the structure in [26], where p is the interpolation order.

An important point concern the gradient computation: the derivative is evaluated on the output signal, at current time for the two components. This imposes a quadrature demodulation in order to separate the envelope and phase of $Z(t)$.

Moreover, the algorithm structure includes tracking capabilities, that are important in case of a non stationnary environment. However its drawback, from the viewpoint of hardware realization and consumption, is that the feedback involves a quadrature demodulation of the output of the transmitter. The suboptimal version below eliminates this need and can function with a simple demodulation or down-conversion.

B. A suboptimal version

It is clear that when the correction algorithm approaches the optimum solution, that is when $\mu_i \to \Delta_i$ or $t_i = t+\mu_i-\Delta_i \to t$, and we have

$$\left.\frac{d\rho(u)}{du}\right|_{t_1}\cos\phi(t_2 \approx \left.\frac{d\rho(u)}{du}\right|_{t}\cos\phi(t) \quad (14)$$

and

$$\rho(t_1)\left.\frac{d\cos\phi(u)}{du}\right|_{t_2} \approx \rho(t)\left.\frac{d\cos\phi(u)}{du}\right|_{t}. \quad (15)$$

This simply suggests that we may substitute the computations at times t_i by the similar computations at time t. These computations can be done using the reference signal at the input and at current time t, rather than requiring the quadrature demodulation of the output. In such a case, the update equations for μ_1, μ_2 become

$$\begin{aligned} \mu_1(n+1) &= \mu_1(n) + \gamma_1(n)\left.\frac{d\rho(u)}{du}\right|_{t}\cos\phi(t)\, e(t), \\ \mu_2(n+1) &= \mu_2(n) + \gamma_2(n)\rho(t)\left.\frac{d\cos\phi(u)}{du}\right|_{t} e(t). \end{aligned} \quad (16)$$

With respect to the original correction algorithm, the new updates only differ by the factor in front of the error term $e(t)$. Therefore, the direction of descent is not affected if these factors have the same sign. Hence, in the domain where these factors have the same sign, the algorithms will converge to the same and true solution. This can give a rough idea of the domain of convergence of the new algorithm. For random processes, this condition shall be considered in mean. We thus consider the domains $\mathcal{S}_i$ of the differences $\mu_i - \Delta_i$ defined by

$$\mathcal{S}_1 : \quad \mathrm{E}\left[\frac{d\rho(u)}{du}\bigg|_{t_1}\cos\phi(t_2)\,\frac{d\rho(u)}{du}\bigg|_{t}\cos\phi(t)\right] \geq 0 \qquad (17)$$

$$\mathcal{S}_2 : \quad \mathrm{E}\left[\rho(t_1)\,\frac{d\cos\phi(u)}{du}\bigg|_{t_2}\rho(t)\,\frac{d\cos\phi(u)}{du}\bigg|_{t}\right] \geq 0 \qquad (18)$$

where the factors exhibit the same sign in statistical mean. Now, as an heuristic guideline, we consider the domain $\mathcal{C}' = \mathcal{C} \cap \mathcal{S}_1 \cap \mathcal{S}_2$, the intersection of the original convergence domains with the $\mathcal{S}_i$. This domain is given in Fig. 6. Compared to Fig. 3, we observe that the domain has been reduced by the approximation. However, the previous condition only gives the region where the two algorithms almost follow the same trajectories. The suboptimal algorithm still converges for delays outside of these regions, but along different paths. The exact determination of the convergence domain of the two coupled non linear equations (16) is a formidable, if not impossible task.

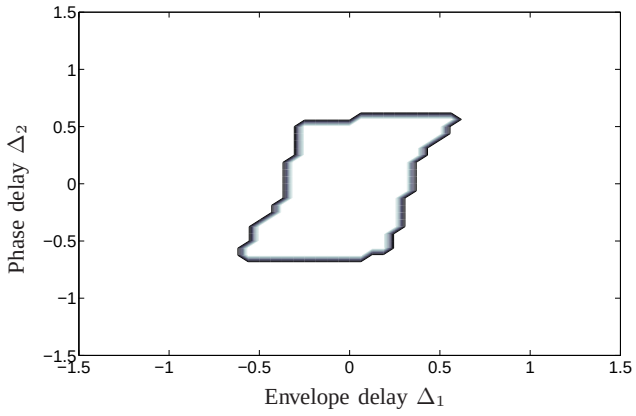


Fig. 6. Domains $\mathcal{C}' = \mathcal{C} \cap \mathcal{S}_1 \cap \mathcal{S}_2$ as the intersection of the domain of convergence $\mathcal{C}$ and the domains $\mathcal{S}_i$ where the gradients evaluated at the system input, at time t, or at its output affected by a delay $\delta_i = \Delta_i - \mu_i$ have the same sign in mean. This gives the indication of the domain where the suboptimal algorithm has the same behavior as the original one.

C. Results and comparisons

Performances have been evaluated using the simulation model presented section II-A and figure Fig. 2. The two algorithms (12) and (16), as well as the interpolator (13) were implemented under ADS. Figure 7 shows the results obtained with the optimum correction algorithm in the case $\Delta_1 = 0.95\ T_s$ and $\Delta_2 = 0.45\ T_s$, for 20 different realizations of a 16QAM sequence.

Table II reports the main figures for the algorithms for different test cases. Both algorithms have the same performances after

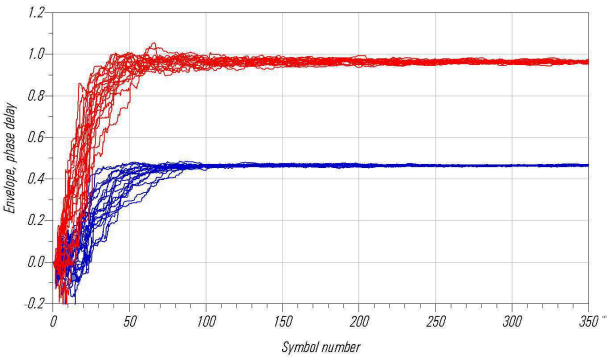


Fig. 7. Results for the optimum correction algorithms in the case $\Delta_1 = 0.95$ and $\Delta_2 = 0.45$, for 20 different realizations of a 16QAM sequence.

the initial convergence. They exhibit a very large improvement for the EVM (falling from 38% in worst case to about 0.4%), as well as for the spectral regrowths, which are reduced to less than 58 dBc. The standard deviation for the estimates of the delays is less than 0.5% of the symbol period. Defining the mean convergence time T_c as the delay to reach the final value with a precision of $0.05T_s$, we obtain that the suboptimal algorithm converges a bit slower than the original one. Convergence speed, bias and variance are also linked to the choice of the adaptation steps γ_1 and γ_2. All the examples in Table II used the same values (normalized to the signal power) $\gamma_1 = \gamma_2 = 0.5$. The fact that there exists a large range of reasonable values for the adaptation steps, with respect to bias, variance and EVM, is discussed in [21].

Finally, we illustrate the tracking capabilities of the procedure in case of variations of the delays. For instance, a drift of some element in the system can induce a slow variation while a commutation between functional modes of the transmitter causes a sudden rupture. The simulation results presented in Fig. 8 show that the solution is also of interest in such situations. It enables to track drifts and plays the role of an autocalibration procedure in case of ruptures. Note that in tracking mode, the algorithm can be updated at a different (much slower) rate than the original sampling rate. It can also be activated only in the case where the drift from the nominal performances excess a preset threshold, or synchronously to a commutation of mode in the system.

IV. Gain and phase identification

The results presented above were obtained without considering the gain and phase offset introduced by the transmitter. Indeed, even with the knowledge of the output power and the global phase shift introduced by the transmitter, it often remains an uncertainty about the gain (0.5 dB for example) and the phase shift of the signal. We first examine the consequences of these uncertainties on the behavior of the algorithms. We show that the algorithms are robust to these errors but that the variance of estimates increases. We then show that it is possible to tackle this problem by the introduction of an identification structure for both the gain and phase offset.

Δ_1	Δ_2	Tc		**EVM (%)**		**Relative Spectrum @**$3.5\ MHz$	
		Opt	Sub Opt	Initial	After	Initial	After
$0.2T_s$	$1.1T_s$	$150T_s$	$190T_s$	38%	0.4%	-30 dBc	-58 dBc
$0.45T_s$	$0.95T_s$	$140T_s$	$160T_s$	25%	0.4%	-32 dBc	-59 dBc
$0.5T_s$	$0.12T_s$	$130T_s$	$150T_s$	20%	0.4%	-35 dBc	-60 dBc
$0.1T_s$	$0.2T_s$	$85T_s$	$110T_s$	5.5%	0.4%	-48 dBc	-68 dBc

TABLE II
ALGORITHMS PERFORMANCES

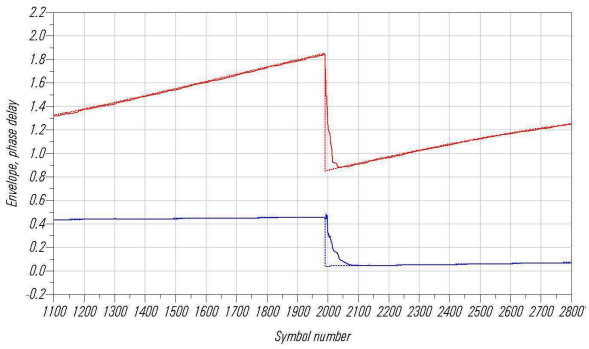


Fig. 8. Illustration of the behavior of the optimum algorithm in case of nonstationarities. This shows that the algorithm can account for slow drifts of the analog system (tracking capabilities) as well as sudden ruptures that may occur in multimodes solutions (autocalibration capabilities).

A. Consequences of uncertainty on gain and phase offset

The direct consequence is that when the input and output of the system are perfectly aligned, the mean square error is not null anymore: if $Z(t)$, the standard output

$$Z(t) = \rho(t-\Delta_1)e^{j\phi(t-\Delta_2)},$$

is affected by a complex gain $K = |K|e^{j\theta}$ the error between the complex envelopes becomes

$$\bar{J}(\Delta_1,\Delta_2) = \mathrm{E}\left[|X(t)-Z(t)|^2\right] = (1+|K|^2)\mathrm{E}\left[\rho(t)^2\right] - 2|K|\mathrm{E}\left[\rho(t)\rho(t-\Delta_1)\cos(\phi(t)-\phi(t-\Delta_2)-\theta)\right], \quad (19)$$

and we obtain $J(0,0) = (1+|K|^2-2|K|\cos(\theta))\mathrm{E}\left[\rho(t)^2\right]$. Hence, even when the signals are aligned, the error is not null. Furthermore, we need to check if the aligned case remains a minimum of the criterion. We note $\epsilon(t) = \phi(t) - \phi(t-\Delta_2)$ the difference process. Its probability density function is known and can be found in [24, 8.82, p. 369] for instance. The phase $\phi(t)$ of a Gaussian process being uniformly distributed on $[0, 2\pi]$, we readily obtain that $\mathrm{E}\left[\epsilon(t)\right] = 0$. As far as the variance is concerned, it is simply given by

$$\mathrm{E}\left[\epsilon(t)^2\right] = 2R_\phi(0) - 2R_\phi(\Delta_2) \quad (20)$$

with $R_\phi(\Delta_2)$ the autocorrelation function of the phase, which can be expressed as a serie [24, 8.81, p. 369].
If we approximate $\epsilon(t)$ and $\rho(t)\rho(t-\Delta_1)$ as uncorrelated variables, an approximation that is reasonable for small delays ($\rho(t)$ and $\phi(t)$ are independent), we have

$$\bar{J}(\Delta_1,\Delta_2) \approx (1+|K|^2)R_\rho(0) - 2|K|R_\rho(\Delta_1)\left(\mathrm{E}\left[\cos(\epsilon)\right]\cos(\theta) + \mathrm{E}\left[\sin(\epsilon)\right]\sin(\theta)\right) \quad (21)$$

with $R_\rho(\tau)$ the autocorrelation function of the envelope $\rho(t)$. Since $\epsilon(t)$ is small, we can expand the expression in ϵ, keep the terms of the development up to the second order and take the expectation using (20). This leads to

$$\bar{J}(\Delta_1,\Delta_2) = (1+|K|^2)R_\rho(0) - 2|K|R_\rho(\Delta_1)\left(1-R_\phi(0)+R_\phi(\Delta_2)\right)\cos(\theta). \quad (22)$$

Consider now Δ_2 fixed and small enough so that $R_\phi(0) - R_\phi(\Delta_2) < 1$. Then, since $R_\rho(\tau)$ is maximum for $\tau = 0$, the criterion is minimum for $\Delta_1 = 0$, provided that $\cos\theta$ is positive. If now Δ_1 is fixed, since $R_\rho(\Delta_1)$ is always positive and $1-(R_\phi(0)-R_\phi(\tau))$ is maximum for $\tau = 0$, we obtain that the criterion is minimum for $\Delta_2 = 0$ if $\cos\theta$ is positive. Therefore, we finally have, as expected, that $J(0,0)$ is a minimum with respect to the two delays. The condition $\cos\theta > 0$ is important and means that the algorithms cannot cope with unknown phase offsets greater than $\pi/2$ (but of course such critical value cannot occur in a correct design). The deformation of the criterion with a phase shift is given in Fig. 9, where we report the criterion with respect to Δ_2, with $\Delta_1 = 0$ fixed, for several values of the phase shift θ. When the phase shift becomes greater than $\pi/2$, then the minimum becomes a maximum.

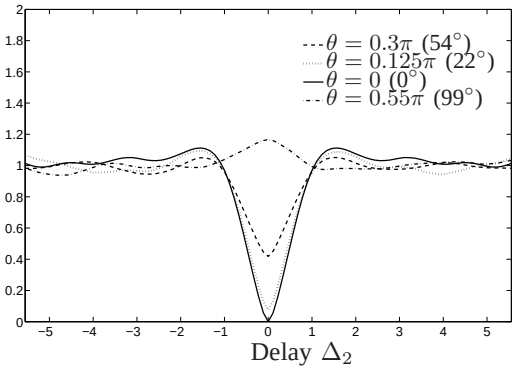


Fig. 9. Criterion with phase mismatches θ. The minimum, whose value increases, remains at $\Delta_2 = 0$, for all phase offsets lower than $\pi/2$.

These results show that the algorithms are robust and can be used even with an imperfect knowledge of the gain, which is important in practical situations. But since this error, whose power converges to a finite value, is used directly in the updates equations of the stochastic algorithms, the convergence is affected by fluctuations. This is illustrated in Fig. 10 where we present a typical result of the optimum algorithm with 0.5 dB of gain mismatch and $10°$ of phase offset.

B. Identification of gain and phase unknowns

Although the algorithms converge in mean to the correct solution, the mean square error is not zero at the optimum and the solution is corrupted by a residual noise. Therefore, it

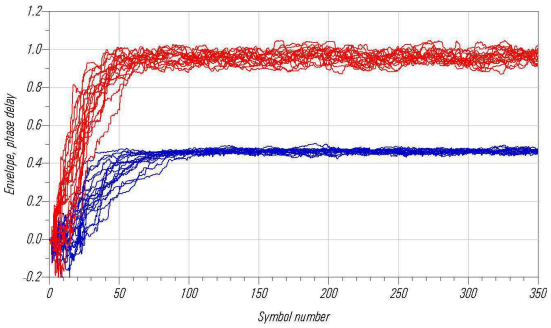


Fig. 10. Delay identification with 0.5 dB of gain mismatch and 10°of phase offset in the case $\Delta_1 = 0.95$ and $\Delta_2 = 0.45$: the algorithm still converges but is corrupted by noise. The fluctuations imply spectral regrowths that can prevent the respect of requirements in performances.

is useful to look for a solution that reduces this residual noise, since it degrades both EVM and spectral performances. Along the lines of the previous developments, the natural solution is to try to identify or correct the unknown gain. In the present case, it is not useful to correct the gain in the transmitter path since after all the receiver has its own equipment for synchronization and gain control, and since other phase shifts and attenuation occur during transmission. Therefore, the gain and phase can be accounted for in the feedback loop, and we can adopt an identification type structure. The complete structure for delays correction together with the account for uncertainty in gain and phase shift is given in Fig. 11.

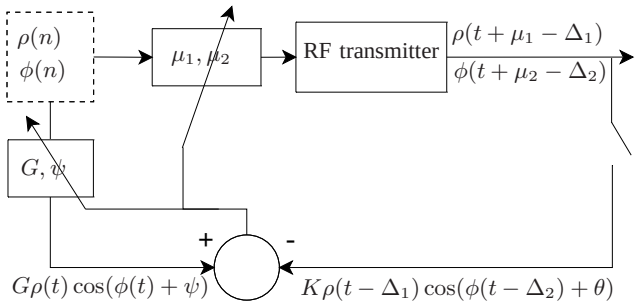


Fig. 11. Complete structure for delays correction with identification of possible gain and phase shift. The same error $e(t)$ is used to correct the delays Δ_1 and Δ_2 introduced in the transmitter and to identify gain variation and phase shift.

As above, the solution can use the in-phase and quadrature components of the output (as in the 'optimum' correction algorithm), or only the in-phase component (as in the suboptimal version). An interest of the identification structure, as shown below, is that the computations on the envelope and phase only involve the digital signals already available at the input. With the notations in Fig. 11, the error is given by

$$e(t) = G\rho(t)\cos(\phi(t)+\psi) - K\rho(t-\Delta_1)\cos(\phi(t-\Delta_2)+\theta). \quad (23)$$

and we look for G and ψ that minimize the mean square error $J(G,\psi) = \mathrm{E}\left[|e(t)|^2\right]$. As before, we derive a gradient solution and implement a stochastic gradient algorithm:

$$\begin{array}{rcl} G(n+1) & = & G(n) - \gamma_3(n)\rho(t)\cos(\phi(t)+\psi(n))e(t) \\ \psi(n+1) & = & \psi(n) - \gamma_4(n)G(n)\rho(t)\sin(\phi(t)+\psi(n))e(t). \end{array} \quad (24)$$

In terms of the complex gain $\tilde{G} = G\exp(j\psi) = G_r + jG_i$ and the complex envelopes, the error can also be expressed as $e(t) = \mathrm{Re}\left(\tilde{G}X(t) - Z(t)\right)$, which leads to the recursions

$$\left\{ \begin{array}{rcl} G_r(n+1) & = & G_r(n) - \gamma_3(n)\rho(t)\cos(\phi(t))e(t), \\ G_i(n+1) & = & G_i(n) + \gamma_4(n)\rho(t)\sin(\phi(t))e(t). \end{array} \right. \quad (25)$$

This formulation is much more interesting for the implementation, since it only involves available signals and does not require evaluations of trigonometric functions of the estimates, as in (24). For completeness, let us also note that the algorithm derived from the complex error $e(t) = \tilde{G}X(t) - Z(t)$, which requires a quadrature demodulation of the output, is

$$\tilde{G}(n+1) = \tilde{G}(n) - \gamma_5 X^*(t)\left(\tilde{G}(n)X(t) - Z(t)\right). \quad (26)$$

C. Results

In the very same conditions as in Fig. 10, that is with 0.5 dB of gain variation and a phase offset of 10°, with $\Delta_1 = 0.95\ T_s$ and $\Delta_2 = 0.45\ T_s$, the introduction of the correction of gain and phase in the feedback loop using (25) leads to results very similar to the original results in Fig. 7 without gain and phase perturbations. These results are quantified in Table III where we consider different scenarios for the gain and phase mismatches. We obtain a large improvement both for the EVM (that falls to about 0.5% near the result without gain and phase mismatch) and for the spectrum at 3.5 MHz, which is attenuated to less than 58 dBc. The only concern is a slower mean convergence which is the price to pay to account for these further mismatches.

V. Conclusions

Polar based transmitter architectures, which are already recognized solutions, are regarded as very promising structures for future realizations. In this work, we considered the critical issue of delays mismatches in polar based transmiter architectures, which degrades the performances, particularly with increasing data rates. We address this problem by digital signal processing. Based on the minimization of the mean square error, we present a baseband algorithm for the correction of delays mismatches. We also propose a suboptimal version that minimizes the realization and implementation costs. We address the fact that the system can also introduce a gain and phase shift. Although it is not necessary to compensate these further mismatches from the system viewpoint, it is useful to account for them in the feedback loop to preserve good performances. This is achieved by the insertion of an identification procedure. Simulation results show a high improvement in the performances, with an EVM below 0.5% and a spectrum level at 3.5 MHz lower than -58 dBc in 30 kHz bandwidth. This demonstrates the real interest of these approaches for the design and optimization of current and future transceivers.

Gain mismatch	phase mismatch	Tc		**EVM** (%)		**Relative Spectrum @**$3.5\ MHz$	
		Opt	Sub Opt	without identification	with identification	without identification	with identification
3 dB	$45°$	$350\ T_s$	$420\ T_s$	23%	0.5%	-40 dBc	-58 dBc
1 dB	$30°$	$300\ T_s$	$400\ T_s$	11%	0.5%	-47 dBc	-58 dBc
0.5 dB	$10°$	$280\ T_s$	$350\ T_s$	3.4%	0.5%	-55 dBc	-59 dBc
0.2 dB	$2°$	$250\ T_s$	$320\ T_s$	0.8%	0.5%	-58 dBc	-60 dBc

TABLE III
ALGORITHMS PERFORMANCES WITH GAIN IDENTIFICATION FOR $\Delta_1 = 0.45\ T_s$ AND $\Delta_2 = 0.95\ T_s$

The exact strategy for the compensation of the delays in a polar architecture has probably to be defined on a case-by-case basis. Since there is no possible useful data transmission during the convergence time, a initial calibration may employ the first correction algorithm, while the suboptimal version can be reserved for refinement and tracking of nonstationarities during operational modes. Of course, the feedback loop can also be used for other compensations, for instance ripples in the frequency response, and the correction algorithm shall cooperate with the compensation of nonlinearities of the power amplifier.

Although it is still possible to further optimize the elements of the transmitter in order to minimize the analog mismatches distorsions, we believe that it is also possible to relax some of the constraints and report a part of the analog complexity to the baseband digital signal processing.

ACKNOWLEDGMENT

The authors are indebted to early readers of this manuscript. The authors also acknowledge O. Venard for several discussions and improvements of the paper.

REFERENCES

[1] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, and N. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, pp. 814–826, 2002.

[2] A. Hietala, "A quad-band 8PSK/GMSK polar transceiver," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 1133–1141, 2006.

[3] J. Groe, "Polar transmitters for wireless communications," *IEEE Communications Magazine*, vol. 45, pp. 58–63, 2007.

[4] J. Groe, "A multimode cellular radio," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, pp. 269–273, 2008.

[5] L. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proceedings of the IRE*, vol. 40, no. 7, pp. 803–806, July 1952.

[6] R. Staszewski, J. Wallberg, S. Rezeq, C.-M. Hung, O. Eliezer, S. Vemulapalli, C. Fernando, K. Maggio, R. Staszewski, N. Barton, M.-C. Lee, P. Cruise, M. Entezari, K. Muhammad, and D. Leipold, "All-digital PLL and transmitter for mobile phones," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2469–2482, Dec. 2005.

[7] D. Rudolph, "Out-of-band emissions of digital transmissions using Kahn EER technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, pp. 1979–1983, 2002.

[8] G. Strasser, B. Lindner, L. Maurer, G. Hueber, and A. Springer, "On the spectral regrowth in polar transmitters," in *Microwave Symposium Digest, 2006. IEEE MTT-S International*, 2006, pp. 781–784.

[9] J.-F. Bercher, A. Diet, C. Berland, G. Baudoin, and M. Villegas, "Montecarlo estimation of time mismatch effect in an OFDM EER architecture," in *Proceedings of the 2004 IEEE Radio and Wireless Conference*, 19-22 Sept. 2004, pp. 283–286.

[10] M. Markovic and H. Modi, "Feasibility of EER transmitters for 3GPP applications," in *IEEE Annual Wireless and Microwave Technology Conference, WAMICON '06*, 2006, pp. 1–2.

[11] W.-F. Loke, M.-W. Chia, and P.-Y. Chee, "Design considerations for multi-band OFDM polar transmitter of UWB system," *Electronics Letters*, vol. 43, 2007.

[12] B. Priyanto, T. Sorensen, O. Jensen, T. Larsen, T. Kolding, and P. Mogensen, "Impact of polar transmitter imperfections on UTRA LTE uplink performance," in *Norchip, 2007*, 2007, pp. 1–4.

[13] W. B. Sander, S. V. Schell, and B. L. Sander, "Polar modulator for multimode cell phones," *Proceedings of the IEEE 2003 Custom Integrated Circuits Conference*, pp. 439–445, 2003.

[14] K. Waheed, R. B. Staszewski, and S. Rezeq, "Curse of digital polar transmission: Precise delay alignment in amplitude and phase modulation paths," in *IEEE International Symposium on Circuits and Systems, ISCAS 2008*, 2008, pp. 3142–3145.

[15] P. Morgan, "Highly integrated transceiver enables high-volume production of GSM/EDGE handsets," *RFdesign*, pp. 36–42, Jul. 2006.

[16] J.-K. Jau and T.-S. Horng, "Linear interpolation scheme for compensation of path delay difference in an envelope elimination and restoration transmitter," in *Proceedings of Asia-Pacific Microwave Conference, 2001. APMC 2001*, vol. 3, 2001, pp. 1072–1075 vol.3.

[17] J. Mártires, C. Borg, and T. Larsen, "Differential delay equalization in a Kahn EER transmitter," in *ST Mobile & Wireless Communications Summit, nr. 15th*, Myconos, Grækenland, Jun. 2006.

[18] C. Zhi and H. Yang, "A new adaptive delay method for wideband Kahn's RF power amplifiers," in *IEEE Tenth International Symposium on Consumer Electronics*, 2006, pp. 1–4.

[19] F. Wang, A. Yang, D. Kimball, L. Larson, and P. Asbeck, "Design of wide-bandwidth envelope-tracking power amplifiers for OFDM applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, pp. 1244–1255, 2005.

[20] J.-F. Bercher and C. Berland, "Adaptive time mismatches identification and correction in polar transmitter architecture," in *2007 European Conference on Wireless Technologies*, 2007, pp. 78–81.

[21] J.-F. Bercher and C. Berland, "Envelope and phase delays correction in an EER radio architecture," *Analog Integrated Circuits and Signal Processing*, vol. 55, pp. 21–35, Apr. 2008.

[22] "3GPP TS 125 213, version 7.0.0, release 7 and 3GPP TS 125 101, version 7.6.0, release 7."

[23] S. C. Cripps, *Advanced Techniques in RF Power Amplifier Design*. Artech House, May 2002.

[24] B. R. Levin, *Fondements théoriques de la radiotechnique statistique*. Moscow: Éditions Mir, 1979, vol. 1.

[25] C. W. Farrow, "A continuous variable digital delay element," in *Proc. IEEE Int. Symp. Circuits Systems*, 1988, pp. 2641–2645.

[26] C. Candan, "An efficient filtering structure for Lagrange interpolation," *IEEE Signal Processing Letters*, vol. 14, no. 1, pp. 17–19, Jan 2007.