Contributions to Statistical Signal Processing with Applications in Biomedical Engineering

Quang Thang Nguyen

N° d'ordre : 2012telb0244

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Bretagne Occidentale

Ecole Doctorale – SICMA

CONTRIBUTIONS TO STATISTICAL SIGNAL PROCESSING WITH APPLICATIONS IN BIOMEDICAL ENGINEERING

Thèse de Doctorat

Mention : STIC – Science et Technologies Information Communication

Présentée par **Quang-Thang NGUYEN**

Département : Signal et Communications

Laboratoire : Lab-STICC Pôle: CID

Directeur de thèse : Dominique PASTOR

Soutenue le 23 Novembre 2012

Jury :

M. Lionel Fillatre – Professeur, I3S (Rapporteur)
M. Alfredo Hernandez – Chargé de recherche (HDR), LTSI INSERM U642 (Rapporteur)
M. Dominique Pastor – Professeur, TELECOM Bretagne (Directeur de thèse)
M. Erwan L'Her – Professeur, LaTIM INSERM U1101 (Examinateur)
M. Lotfi Senhadji – Professeur, LTSI INSERM U642 (Examinateur)
M. Emanuel Radoi – Professeur, UBO/Lab-STICC CNRS UMR 6285 (Examinateur)
M. Ronan Fablet – Maître de conférence (HDR), TELECOM Bretagne (Examinateur)
M. François Lellouche – Professeur, Université Laval (Québec –Canada) (Invité)

Contributions to Statistical Signal Processing with applications in biomedical engineering

Quang-Thang NGUYEN

April 17, 2013

Remerciements

Ces travaux de thèse ont été réalisés au sein du département Signal et Communications de TELECOM Bretagne. La thèse est financée par l'Institut MINES-TELECOM.

Je tiens tout d'abord à remercier Monsieur Lotfi Senhadji, Professeur au LTSI (Laboratoire Traitement du Signal et de l'Image), pour avoir accepté de participer et de présider ce jury de thèse. Je remercie également les rapporteurs, Monsieur Lionel Fillatre, Professeur à l'I3S (Laboratoire d'Informatique, Signaux et Systèmes de Sophia-Antipolis) et Monsieur Alfredo Hernandez, Chargé de recherche (HDR) au LTSI (Laboratoire Traitement du Signal et de l'Image) pour leur lecture de ce manuscrit ainsi que pour leurs points de vue et leurs remarques pertinentes sur cette thèse. Je suis reconnaissant à Monsieur Emanuel Radoi, Professeur à l'UBO/Lab-STICC, d'avoir accepté d'examiner ces travaux.

Des remerciements particuliers vont au Monsieur François Lellouche, Professeur à l'Université de Laval (Québec, Canada) pour m'avoir mis à disposition une grande base de données et réalisé un travail important d'analyse clinique sans lesquels ces travaux n'auraient pu être validés.

Je souhaite exprimer toute ma gratitude et mes remerciements à mon directeur de thèse, Monsieur Dominique Pastor, Professeur à TELECOM Bretagne, pour m'avoir donné la possibilité d'entreprendre mes travaux de doctorat, ainsi que pour de sa disponibilité et son enthousiasme. Travailler avec lui a été à la fois très agréable et très enrichissant scientifiquement.

Mes grands remerciements vont aussi à mes encadrants, Monsieur Ronan Fablet, Maître de Conférence (HDR) à TELECOM Bretagne et Monsieur Erwan L'Her, Professeur au LaTIM/Hôpital Cavale Blanche pour m'avoir guidé sur ces thématiques de recherche. Merci pour vos conseils, votre disponibilité et votre assistance.

Ces trois années de thèse n'auraient pas été particulièrement réussies sans l'ensemble des permanents, des doctorants et des stagiaires que j'ai eu l'occasion de rencontrer au département Signal et Communications de TELECOM Bretagne et qui ont contribué à créer une ambiance de travail très agréable. Merci pour vos conseils et vos amitiés qui ont fait ces trois années des moments inoubliables.

Un énorme merci va à mes amis qui m'ont soutenu.

Pour finir, je tiens à exprimer ma très profonde gratitude et mes très grands remerciements envers ma famille, mes parents, ma sœur et en particulier, à Minh-Huong Nguyen pour leur amour, leur confiance, leur soutien et leurs encouragements.

Merci à vous !

Sommaire

Problématiques

Ces dernières années, le domaine de l'ingéniérie biomédicale a subi une croissance explosive. De nombreux problèmes ont été présentés comme des défis qui nécessitent les efforts conjoints de chercheurs et d'ingénieurs de différentes disciplines afin d'offrir une solution. Ces études multidisciplinaires sont devenus de plus en plus importantes non seulement pour relever ces défis, mais aussi enrichir les compétences de chaque domaine. Au sein du laboratoire Lab-STICC/CID/TOMS du département Signal et Communications de TELECOM Bretagne, de nombreuses techniques de la théorie de l'information et du traitement statistique du signal — en particulier, de parcimonie, statistiques robustes, détection statistique/classification/estimation — ont été développées. Le but de cette thèse de doctorat est de contribuer à ces expertises et savoir-faires à travers des applications en ingéniérie biomédicale. En combinant les compétences internes du laboratoire et celles des experts du domaine d'application, notre tentative est non seulement de résoudre des problèmes dans ce domaine d'application, mais aussi d'améliorer la base théorique, la méthodologie et le transfert de connaissances vers d'autres champs d'application possibles, au-delà de l'application biomédicale. Dans ces travaux de thèse, la décision statistique sera considérée à travers deux applications concrètes: la détection des hotspots à l'interface de protéines et la surveillance automatique des systèmes de ventilation mécanique en médecine d'urgence.

La décision statistique, qui est au coeur de ces travaux, est un de problèmes majeurs du traitement statistique du signal. Supposons que l'on ait un ensemble d'observations dont les distributions dépendent de certaines hypothèses, une décision statistique est une fonction qui associe chaque observation donnée à une décision. Cette décision n'est qu'un choix entre les différentes hypothèses. En fonction de ce qu'on connaît sur les distributions conditionnelles des observations sachant chacune des hypothèses, différentes approches paramétriques ou non-paramétriques sont proposées. Si ces distributions sont connues a priori, un ensemble de méthodes paramétriques visent à trouver une solution optimale en minimisant la probabilité d'erreurs, en maximisant la probabilité de détection sous contrainte d'une probabilité de fausse-alarme bornée, ou encore min-

imiser la probabilité d'erreur dans le cas le plus défavorable. Parmi les approches principales, nous pouvons citer les méthodes de Bayes, de Neyman-Pearson, minimax, etc. Inversement, si aucune information sur ces distributions conditionnelles n'est donnée, des approches non-paramétriques sont étudiées. La différence entre les approches repose sur la façon dont ce manque d'information est compensé. Au cas où une base suffisamment grande et descriptive de données est disponible, les machines d'apprentissage proposent d'estimer ces distributions à travers ces données disponibles. Dans nombreux cas, la construction d'une telle base de donnée est très couteuse, voire impossible. Si le modèle d'observation est partiellement connu, par exemple, lorsque la distribution du bruit est connue mais pas celle du signal, des tests robustes d'hypothèses, notamment, le Random Distortion Testing (ou RDT), peuvent être considérés comme de bons candidats. Les deux applications biomédicales étudiées dans ces travaux représentent deux cas typiques de ces deux approches non-paramétriques.

Aussi, cette thèse est divisée en deux parties. Dans la première partie, la détection de hotspots, qui sont des résidus critiques d'interaction de protéines, est traité en utilisant la méthode de machine apprentissage. Pour cette application, nous avons proposé une nouvelle famille de descripteurs basés uniquement sur les caractéristiques fréquentielles de séquence de protéine. En utilisant ces descripteurs à travers une forêt aléatoire, nous arrivons à bien identifier les hotspots. Dans la deuxième partie, le Random Distortion Testing est étudié à travers l'application de surveillance du système de ventilation artificielle. Les contributions majeures dans ce domaine incluent : trois problèmes de détection dans le cadre de RDT, l'extension du RDT sous forme RDT dual séquentiel (RDT-DS) et finalement, la détection automatique d'anomalie en ventilation artificielle. À notre connaissance, la détection automatique d'autoPEEP que nous proposons est le premier système au monde qui permet de détecter ce type fréquent d'anomalie.

Première partie : machine d'apprentissage et détection de hotspots

La première partie, consacrée à la décision statistique dans le cadre des machines d'apprentissage et l'application de détection de hotspots d'interaction des protéines, se compose de deux chapitres.

Machine d'apprentissage

Dans le chapitre 1, les notions principales concernant les approches sur les machines d'apprentissage, en particulier les forêts aléatoires, sont résumées. Pour une machine d'apprentissage, nous disposons d'une base de données constituée d'observations et

de labels. Dans la phase d'apprentissage, cette base de données est utilisée pour estimer une fonction de décision. En fonction de ce que nous connaissons sur les labels, nous avons différents types d'apprentissage: supervisé, non-supervisé, semi-supervisé, présence/absence, faiblement supervisé. Dans ces travaux, nous nous intéressons à l'approche supervisée qui suppose que le label est connu pour chaque échantillon de la base d'apprentissage. Pour notre application de détection de hotspots, les forêts aléatoires sont utilisées grâce à leurs avantages par rapport à d'autres méthodes de l'état de l'art, notamment, leur performance de classification et le petit nombre de paramètres à régler. Une forêt aléatoire est une technique d'apprentissage automatique qui est composée de plusieurs arbres de décision. Chaque arbre de décision représente une segmentation de l'espace d'observation (ou espace de descripteurs) en régions homogènes. Chaque région homogène est associée à un label. L'homogénéité est définie à l'aide d'un critère entropique. En pratique, la notion d'entropie de Shannon et la diversité de Gini sont souvent utilisées. En combinant plusieurs arbres de décision par vote majoritaire, la forêt aléatoire vise à donner une segmentation optimale.

Application à la détection des hotspots

Au chapitre 2, l'application à la détection de hotspots d'interaction des protéines est présentée. Comprendre la structure et la fonction biologique des protéines, les blocs de construction de base de tous les organismes vivants, est l'un des sujets les plus importants de la biologie. En général, les biochimistes distinguent les structures de protéines en quatre niveaux différents. Au premier niveau, une protéine présente simplement une séquence d'acides aminés. On compte 20 différents acides aminés dans la nature. Au deuxième niveau, cette séquence se replie localement sur elle-même pour former des structures locales, comme des hélices alpha ou des feuillets beta, qui s'appellent structures secondaires. La structure au troisième niveau représente la structure 3D complète de la protéine. Pour les protéines complexes qui contiennent plus d'une chaine polypeptidique, les chercheurs distinguent aussi la structure au quatrième niveau qui décrit la façon dont ces sous-unités interagissent et s'imbriquent en 3D.

Les études de l'état-de-l'art de biologie ont conclu qu'une protéine fonctionne en formant une structure 3D active et interagissant avec son partenaire via une interface de forme géométrique complémentaire. La distribution de l'énergie de liaison sur cette interface n'est pas uniforme. Il y a des résidus qui contribuent plus d'énergie que les autres. Ces résidus s'appellent des hotspots d'interaction de protéines. Une fois ces résidus mutés, l'interaction peut être déstabilisée et la fonction de protéine est altérée. Aussi, l'identification des hotspots est la clé pour mieux comprendre les fonctions de protéines.

Expérimentalement, les hotspots sont identifiés par mutagénèse. Mais cette tech-

nique est coûteuse en temps et en moyens. Donc, il faut des méthodes calculatoires alternatives. Parmi les méthodes de l'état de l'art, les méthodes les plus performantes nécessitent l'information sur la structure 3D de la protéine. Mais, sachant que toute l'information sur la structure et la fonction de protéine s'écrit dans sa séquence primaire d'acides aminés, il est donc possible d'avoir une détection de hotspots basée uniquement sur la structure 1D primaire de protéine. À la recherche d'une telle méthode, I. Cosic a trouvé l'existence d'une fréquence caractéristique pour chaque famille fonctionnelle de protéines. Cette fréquence correspond à une périodicité dans la distribution d'énergie d'électrons de covalence le long de la chaine polypeptidique qui caractérise la façon dont une protéine peut reconnaitre et interagir avec son partenaire. Ce modèle physico-mathématique s'appelle le Modèle de Reconnaissance par Résonance (RRM pour Resonant Recognition Model). En termes de RRM, les hotspots sont les résidus les plus affectés par une modification de la fréquence caractéristique. Motivées par cette notion, des méthodes reposant sur l'identification de la fréquence caractéristique ont été proposées. Mais ces méthodes ne sont pas pratiques pour les nouvelles protéines séquencées de famille inconnue. La fréquence caractéristique ne peut pas alors être calculée. En plus, un changement dans le domaine fréquentiel va affecter tous les échantillons dans le domaine "temporel".

Pour éviter ces inconvénients, nous proposons une nouvelle approche qui combine la mutagénèse et le modèle RRM. En particulier, nous utilisons une mutation symbolique pour analyser les caractéristiques fréquentielles de la séquence de protéine et en déduire des descripteurs. Une classification par forêt aléatoire est alors utilisée pour identifier les hotspots. L'avantage majeur de cette approche est que nous restons toujours dans le domaine fréquentiel et ne revenons jamais par une transformation inverse au domaine "temporel" pour localiser les hotspots. Figure 2.2 présente le schéma principe de notre méthode. Pour un résidu d'intérêt, nous le remplaçons symboliquement par un résidu alanine. La séquence originale et la séquence mutée sont analysées et comparées pour extraire des descripteurs. Dans ces travaux, trois types de descripteurs ont été considérés: le changement d'énergie global, le changement d'énergie dans les sous-bandes et le changement des pics. Nous pouvons considérer ces trois descripteurs comme une analyse multi-résolution du spectre de séquence de protéine. Cette analyse nous permet de capitaliser sur les caractéristiques fréquentielles potentiellement liées à l'identification de hotspots.

L'évaluation a été faite sur une base de 221 résidus. Les forêts aléatoires sont utilisées pour la classification. Et les résultats sont rapportés dans les tableaux 2.3, 2.4 et sur la figure 2.3. Nous pouvons constater que les descripteurs 1D que nous proposons fonctionnent mieux que les descripteurs 3D de l'état de l'art et la performance de ces descripteurs est supérieure à celle donnée dans la littérature. La combinaison avec les descripteurs 3D améliore encore la performance d'identification. En utilisant la forêt

aléatoire, cette combinaison permet une identification avec une exactitude de 82% et une précision de 80%.

Ces résultats de classification montrent que les descripteurs proposés sont pertinents pour l'identification des hotspots. Ces résultats confirment aussi le fait que les fonctions de protéine sont codées dans leurs séquences primaires d'acides aminés. Mais comment ce codage est-il fait? Pour répondre à cette question, il faut étudier et approfondir le sens physique des descripteurs fréquentiels proposés. La simplicité de la méthode nous permettrait de faire une telle analyse sur une plus grande base de données que celle dont on dispose actuellement. En perspective, la combinaison de cette méthode avec d'autres approches de machine d'apprentissage comme le SVM ou l'apprentissage non-supervisé/semi-supervisé peut être considéré. D'autres descripteurs peuvent aussi être proposés.

Deuxième partie : statistiques semi-paramétriques robustes et application à la surveillance des signaux respiratoires en médecine d'urgence

Dans la première partie, nous avons utilisé les machines d'apprentissage. L'inconvénient des machines d'apprentissage que nous avons considérées est de reposer sur la nécessité d'une base de données suffisamment grande et descriptive. En plus, les machines d'apprentissage sont limitées en termes de performance qu'elles peuvent garantir. Cette performance est conditionnée par la base de données qu'on dispose. Pour aller au delà de cette limitation et, donc, échapper aux contraintes sur la base de données, il faut employer un modèle d'observation suffisamment générique pour englober un grand nombre d'applications. Dans la deuxième partie, un tel modèle est considéré pour des problèmes de détection dans le cadre de Random Distortion Testing (RDT) et l'application de surveillance des systèmes de ventilation mécanique. Cette partie se compose de trois chapitres.

Random Distortion Testing (RDT) et RDT dual séquentiel

Au chapitre 3, les notions de base du RDT sont résumées. Le modèle générique suivant est employé pour affranchir la nécessité d'une base de données quand celle-ci n'est pas disponible dans les applications.

$$\begin{cases} \text{Observation:} & \mathbf{Y} = \boldsymbol{\Theta} + \mathbf{X} \begin{cases} \mathbf{X} \sim \mathcal{N}(0, \mathrm{C}) \\ \boldsymbol{\Theta} \text{ et } \mathbf{X} \text{ sont independants} \end{cases} \\ \text{Événement testé } (h_0): & \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant \tau \\ \text{Événement alternatif } (h_1): & \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau \end{cases}$$

En fait, nous considérons l'observation d'un signal aléatoire de distribution inconnue dans un bruit Gaussien, additif et indépendant. Le problème est de tester si une réalisation du signal est égal ou non à une référence connu $\boldsymbol{\theta}_0$. Ce problème est motivé par des applications pratiques comme le radar, le sonar ou la surveillance de signaux. Dans le cas où le signal est déterministe inconnu et le nombre d'échantillons est suffisamment grand, on peut utiliser les tests classiques comme de Neyman-Pearson [Neyman and Pearson, 1928], de Rao [Rao, 1948] ou de Wald [Wald, 1943] pour avoir une bonne performance. Dans le cas général d'un signal aléatoire de distribution inconnue, le problème est ouvert. Dans [Pastor and Nguyen, 2012a], on utilise l'invariance du bruit gaussien pour proposer le cadre théorique du Random Distortion Testing (ou RDT). En pratique, tester sur le signal bruité égal ou non le modèle $\boldsymbol{\theta}_0$ est trop sévère, voire impossible, à cause de perturbations dues à l'environnement, même en absence du bruit. En plus, dans certaines applications, nous ne sommes intéressés qu'à détecter des distorsions suffisamment grandes et non pas les distorsions de faibles amplitudes. D'où l'intérêt d'avoir une valeur de tolérance τ. Le problème revient alors à tester si la distance entre le signal et le modèle est supérieure ou inférieure à cette valeur de tolérance τ. La norme utilisée est la norme de Mahalanobis pour compenser la variation introduite par la matrice de covariance du bruit. Le problème est invariant par rapport aux sphères, pour la norme de Mahalanobis, que sont les orbites du groupe qui laisse le bruit invariant. Ainsi, si la réalisation du signal tombe sur une sphère de rayon ρ autour de la référence $\boldsymbol{\theta}_0$, le problème est inchangé. Le test optimal proposé par le RDT est alors le suivant:

$$\mathcal{T}_{\lambda_\gamma(\tau)}\left(\mathbf{Y}(\omega)\right) = \begin{cases} 1 & \text{si} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| > \lambda_\gamma(\tau) \\ 0 & \text{si} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| \leqslant \lambda_\gamma(\tau) \end{cases}$$

Le seuil optimal $\lambda_\gamma(\tau)$ est calculé à l'aide de l'équation $1 - \gamma = F_{\chi^2_d(\tau^2)}(\eta^2)$ où γ est la valeur maximale de la probabilité de fausse-alarme et $F_{\chi^2_d(\tau^2)}(.)$ est la fonction de répartition de la loi Chi-deux non-centré avec d degrées de liberté et de paramètre de non-centralité ρ^2. Le test RDT proposé est MCCP (pour *maximal constant conditional power*). En fait, ce test a une puissance constante sur chaque sphère et cette puissance constante est supérieure sur toute sphère de rayon $\rho > \tau$ à celle de n'importe quel autre test de même taille et de puissance constante sur la même sphère. La borne inférieure de la puissance du test est aussi donnée (cf. Eq. 3.19).

A la fin de ce chapitre, l'application du RDT à la détection du signal, un problème classique en traitement du signal, est étudiée. Les résultats montrent que le test RDT est résistant aux variations du signal et à l'imperfection du modèle d'observation. En vertu de ce qui précède, nous pouvons considérer le RDT comme une solution alternative aux approches classiques, par exemple l'approche de Neyman-Pearson, dans le cas où il existe une incertitude sur le modèle du signal utilisé.

Trois problèmes de détection dans le cadre du RDT

Au chapitre 4, motivé par notre application à la surveillance des systèmes de ventilation artificielle où le signal est temporel et unidimensionnel, nous considérons trois problèmes de détection dans le cadre de RDT. Le problème [**Dev.**] est la détection de déviations du signal à des instants critiques, comme la détection d'autoPEEP en ventilation artificielle qui sera décrite ci-desous. Ce type de problème de détection est l'occasion d'introduire le RDT dual et le RDT séquentiel. Le problème [**Chg.**] est celui de la détection de changement de phase. Ce problème se rencontre dans des applications comme la détection d'ondes dans le signal ECG ou la segmentation de phases respiratoires en ventilation artificielle. Pour le troisième problème [**Dis.**], nous nous intéressons à détecter une distorsion de distribution inconnue d'un signal dans un intervalle du temps, comme la détection d'efforts inefficaces en ventilation artificielle. Ces trois problèmes sont classiques en traitement du signal. Dans ce chapitre, nous allons les reformuler et les résoudre dans un nouveau cadre de test robuste d'hypothèse RDT.

Pour le problème [**Dev.**], nous observons un signal temporel bruité et nous nous intéressons à détecter la déviation du signal à un instant critique t_c. En prenant en compte un voisinage de l'instant t_c, nous avons l'observation sous forme vectorielle. Nous pouvons factoriser ce vecteur sous la forme $\mathbf{p}\Theta(t_c)$ où $\Theta(t)$ est le signal. Ce vecteur $\mathbf{p}$ représente la forme locale du signal autour de l'instant critique. En faisant une projection de l'observation sur la direction engendrée par le vecteur de forme locale $\mathbf{p}$, nous avons une nouvelle observation avec un niveau de bruit plus faible. La décision proposée est alors celle donnée par le test RDT.

Pour K instants critiques consécutifs, nous supposons que le signal varie lentement et que la référence reste inchangée. Pour chaque instant t_k de décision, nous observons le même signal en présence d'une réalisation différente de bruit. La détection peut se faire par une analyse séquentielle qui repose sur deux éléments principaux: Premièrement, il faut un nouveau test, le test RDT dual, basé sur deux seuils qui nous renvoie que des décisions fiables dans le sens où les deux seuils sont calculés de manière pour garantir une très faible valeur de la probabilité de fausse-alarme et une très forte valeur de la probabilité de détection. Le principe est que nous ne prenons pas de décision lorque l'obserbation n'est pas suffisamment pertinente pour prendre une décision avec une

faible probabilité d'erreur. Dans ce cas, la décision est retardée pour intégrer de nouvelles observations qui permettront de prendre la décision. La méthode la plus simple pour intégrer ces nouvelles observations consiste à moyenner celles-ci. Lorsque le nombre d'observations augmente, l'écart-type du bruit intégré diminue et tend vers 0. Nous pouvons alors montrer que, quand la variance du bruit tend vers 0, le test RDT dual tend vers un test dont la probabilité de fausse-alarme tend vers 0 et la probabilité de détection tend vers 1. Le test RDT séquentiel effectue son analyse séquentielle comme suit: nous commençons par tenter une décision sur la première observation à l'aide d'un test RDT dual. Si nous ne parvenons pas à une décision sur cette première observation, nous acquerrons la deuxième observation que nous moyennons avec la première. Si nous ne pouvons toujours pas prendre une décision à partir de ces observations moyennées, nous intégrons une troisième observation et continuons à procéder ainsi de proche en proche. Le processus s'arrête quand une décision a été prise. Si une décision ne peut être prise par un test RDT dual à la M-ième observation, nous arrêtons le processus par une décision forte qui consiste à n'utiliser que le seuil le plus fort, ce qui permet de garantir la probabilité de fausse-alarme. Cette décision forte est un test RDT classique. L'optimalité est alors celle du RDT classique. Sur la figure 4.1, nous voyons que les deux seuils tendent vers la même valeur de tolérance τ et la région de non-décision tend vers 0 quand le nombre d'observations tend vers l'infini.

Pour le problème [**Chg.**], nous nous intéressons à localiser les instants où ont lieu les changements de phase. Supposons que le signal soit régulier dans chaque phase et que le changement de l'état interne de la source du signal génère des irrégularités dans signal. A titre d'exemple, pour un signal de débit respiratoire, nous nous intéressons à identifier les fins d'inspirations et les fins d'expirations. Notre méthode proposée est alors composée de deux étapes consécutives : en première lieu, une transformation parcimonieuse, par exemple la transformation d'ondelettes, est utilisée pour mettre en évidence les irrégularités dans le signal observé. Un seuillage est ensuite effectuée pour détecter les pics de la transformée en ondelettes en utilisant le test RDT avec une valeur de tolérance égale au seuil universel. On peut noter que le seuil universel peut être considéré comme la valeur absolue maximale du bruit quand le nombre d'échantillons tend vers l'infini. Les pics retenus correspondent aux irrégularités du signal ou alors aux instants de change de phase. Figure 4.4 présente le résultat de la détection de changement de phase respiratoire sur le signal de débit donné par une telle méthode. L'idée d'utiliser le seuil universel comme la valeur de tolérance nous permet d'éviter les fausse-alarmes sur ces points.

Dans le problème [**Dis.**], nous nous intéressons à la détection de distorsions d'un signal par rapport à une référence dans un intervalle du temps. En prenant en compte le vecteur d'observation qui contient tous les échantillons du signal observé dans l'intervalle d'intérêt, le problème revient exactement au cas général multidimension-

nel du RDT et la décision est donnée par le test MCCP de la théorie.

Ces trois problèmes se rencontrent tels quels dans l'application à la surveillance du système de ventilation artificielle étudiée au chapitre 5.

Application à la ventilation artificielle

La ventilation artificielle est pour le but d'assister ou remplacer la respiration spontanée de patient souffrant de pathologies respiratoires. Cette technique est utilisée dans les services d'urgence, de réanimation ou à domicile. Malheureusement, l'interaction entre le ventilateur et le patient est imparfaite. Il y a souvent des anomalies, par exemple l'autoPEEP ou l'asynchronie qui génèrent l'assistance incomplète ou l'augmentation de l'effort respiratoire. Les études de l'état de l'art ont montré que les courbes respiratoires disponibles chez les ventilateurs actuels (comme les courbes de débit, de volume et de pression) nous donnent l'information pertinente pour détecter les anomalies. Parmi les études de l'état de l'art, des méthodes visent à identifier automatiquement les efforts inefficaces, les doubles déclenchements. Malheureusement, à notre connaissance, la détection automatique de l'autoPEEP et d'autres types d'asynchronie n'a pas encore été étudiée. C'est le but de notre application.

L'AutoPEEP (pour *auto-positive end expiratory pressure*) est un phénomène où la pression à la fin d'expiration reste positive. Cette anomalie est causée par un temps expiratoire insuffisant. Sur la courbe de débit, les AutoPEEPs peuvent être identifiés par les non-retours à zéro du signal à la fin d'expiration. Cette détection visuelle est simple mais elle nécessite la présence d'un clinicien au côté du patient. Cela n'est pas toujours possible en pratique. Il est donc nécessaire d'avoir une détection automatique pour optimiser la tâche de clinicien et pour pouvoir employer une surveillance en continue. La figure 5.2 décrit le schéma principe de la plateforme de détection proposée dans ce mémoire. À l'entrée, nous avons une acquisition et une conversion de données pour former le vecteur d'observation. Le détecteur d'autoPEEP, le coeur du système, est effectivement l'application du problème **[Dev.]** que nous avons étudié au chapitre 4. Deux détecteurs d'autoPEEP sont proposés : un basé sur le RDT classique et l'autre basé sur le RDT séquentiel. Le détecteur de changement de phases respiratoires est l'application directe du problème **[Chg.]**.

Pour la détection, il faut aussi que certains paramètres soient estimés, notamment le vecteur de forme $\mathbf{p}$ et l'écart type du bruit. Sachant que le signal de débit dans la phase expiratoire s'exprime sous forme exponentielle, le vecteur de forme locale $\mathbf{p}$ peut être estimé par une régression. Plusieurs régressions peuvent aussi être agrées pour une meilleure estimations. Pour l'écart type du bruit, les estimateurs de l'état de l'art comme le MAD (pour *median absolute deviation*) ou le DATE (pour *d-dimensional adaptive trimming estimator*). Dans notre application, le MAD et le DATE donne des

résultats similaires mais le MAD fonctionne plus vite.

Les évaluations ont été faites en trois différents niveaux: par simulations, par émulations avec un poumon artificiel programmable, et finalement, par l'analyse sur les courbes cliniques. Les résultats des simulations sont illustrées par les figures 5.9 et 5.10. Les performances de la détection dans les différentes configurations d'émulation sont présentés au tableau 5.1. L'analyse rétrospective sur les courbes cliniques enregistrées sur les patients anonymes au Service d'Urgence du Centre Hospitalier Universitaire Centre Hospitalier Universitaire de Brest (France) et à l'Institut Universitaire de Cardiologie et de Pneumologie de Québec (Canada) a montré que, sur une base de 15 patients et environ 1998 cycles, notre plateforme donne une très bonne détection avec une exactitude de 93% et un rappel de 90%. Ces valeurs sont très pertinentes pour les applications cliniques. La vérité terrain est donnée par une analyse clinique d'une équipe d'experts du domaine.

La plateforme proposée peut aussi être étendue à la détection d'asynchronies. En fonction de leur nature, nous distinguons deux types d'asynchronies. Les asynchronies du premier type sont celles liées au temps de déclenchement, par exemple, les cycles courts (cf. Figure 5.14(a)), les inspirations prolongées (cf. Figure 5.14(b)) et les doubles déclenchements (cf. Figure 5.15). La détection de ces types d'asynchronie est une application de la détection de changement de phases respiratoires, ou problème [**Chg.**], que nous avons traité au chapitre 4. Les asynchronies du deuxième type sont celles liées aux distorsions de la forme d'onde. La détection est, en fait, l'application directe du problème [**Dis.**] — la détection de distorsion dans un intervalle du temps. Dans ces cas, nous considérons les efforts inefficaces pendant l'expiration (cf. Figure 5.16) à titre d'exemple. L'évaluation sur les signaux synthétisés a montré que, pour une valeur maximale de taux de fausse-alarme de $\gamma = 0.01$, notre algorithme arrive à un taux de détection de 90%. Les courbes de performance sont tracées sur le Figure 5.18.

Pour aller au delà des limitations de l'évaluation sur les données clinique, un simulateur virtuel de ventilation mécanique (cf. Figure 5.20) a été conçu. En simulant plusieurs types de patients avec différentes conditions de santé, ce simulateur nous permet d'établir une évaluation exhaustive des algorithmes proposés sur différents cas, incluant des cas rarement vus en pratique. Il permet aussi des tests en boucle fermée qui est strictement réglementé sur les vrais patients.

Etant donné les différents détecteurs d'anomalie, un système de surveillance à distance et en temps continu est aussi proposé (cf. Figure 5.19). Au chevet des patients, un moyen d'acquisition de données permet d'envoyer le signal à un serveur central équipé. La détection des anomalies peut se faire soit sur le serveur central, soit au chevet du patient. Le signal d'alarme sera envoyé au clinicien correspondant via différents types d'équipement dont il dispose. Les informations envoyées sont adaptées à la qualité de la connexion. Par exemple, si le clinicien ne dispose que d'un téléphone portable, seul

le signal d'alerte lui sera envoyé. S'il dispose d'un lien plus performant, par exemple un ordinateur portable avec une connexion Wifi, les signaux respiratoires (débit, pression, volume, etc) pourront lui être envoyés en plus des résultats de détection pour lui permettre une analyse plus approfondie. Ce système permet l'intervention plus rapide et l'optimisation du soin des cliniciens. Vu le nombre de ventilateurs à domicile en France et aux États Unis, nous voyons que ce type du système peut aider à améliorer la qualité de vie de patients.

Conclusions et perspectives

Dans ces travaux de thèse, nous avons considéré deux approches non-paramétriques avec deux applications typiques en ingéniérie biomédicale. Pour les machines d'apprentissage, nous avons la possibilité d'utiliser plusieurs types de descripteurs. Par contre, il nous faut une base de données suffisamment grande et descriptive. Ces approches présentent aussi une difficulté dans l'interprétation des résultats. Le RDT nécessite un modèle d'observation et la variance du bruit mais cette variance peut être estimée en pratique. Il n'y a pas d'apprentissage, pas de base de données. Ainsi, on peut considérer le RDT comme une approche semi-paramétrique car il nous donne l'optimalité et la robustesse par rapport aux variations du signal et l'imperfection du modèle d'observation. Il nous permet aussi de prendre l'expérience du spécialiste du domaine en compte. Par exemple, pour notre application de ventilation artificielle, l'expérience du clinicien est pris en compte via la valeur de tolérance τ et le taux de fausse-alarme γ.

Cette thèse a été financée par l'Institut MINES-TELECOM. Certains aspects théoriques et commerciaux de l'application de surveillance du système de ventilation artificielle seront étudiés dans le cadre du projet CURVEX, financé par l'Institut MINES-TELECOM et la Région Bretagne dans le but de monter une "Start-Up".

Contents

Abbreviations

ASA	Accessible Surface Area
ASEdb	Alanine Scanning Energetics database
ASM	Alanine Scanning Mutagenesis
AutoPEEP	Auto-Positive End Expiratory Pressure
BID	Binding Interface Database
DATE	(noise estimator) d-Dimensional Adaptive Trimming Estimator
DDT	Deterministic Distortion Testing
DSP	Digital Signal Processing
ECG	ElectroCardioGraphy
EIIP	Electron-Ion Interaction pseudo-Potential
EMC	ElectroMagnetic Compatibility
FFT	Fast Fourier Transform
FGF	(protein family) Fibroblast Growth Factor
GLRT	Generalized Likelihood Ratio Test
HOTPOINT	(hotspot decriptors) Solvent accessibility and pair potentials
IC	Ionization Constant
IEE	Ineffective Effort during Expiration
KFC	Knowledge-based FADE and Contacts
MAD	(noise estimator) Median Absolute Deviation
PDB	Protein Data Bank
RDT	Random Distortion Testing
RF	Random Forests
Robetta	(hotspot descriptor) Computational binding free energy change
RRM	Resonant Recognition Model
SNR	Signal-to-Noise Ratio
SONAR	SOund Navigation and Ranging
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
UMP	Uniformly Most Powerful

List of Figures

List of Tables

Abstract

This PhD thesis presents some contributions to Statistical Signal Processing with applications in biomedical engineering. The thesis is separated into two parts.

In the first part, the detection of protein interface hotspots — the residues that play the most important role in protein interaction — is considered in the Machine Learning framework. The Random Forests is used as the classifier. A new family of protein hotspot descriptors is also introduced. These descriptors are based exclusively on the primary one-dimensional amino acid sequence. No information on the three dimensional structure of the protein or the complex is required. These descriptors, capturing the protein frequency characteristics, make it possible to get an insight into how the protein primary sequence can determine its higher structure and its function.

In the second part, the RDT (Random Distortion Testing) robust hypothesis testing is considered. Its application to signal detection is shown to be resilient to model mismatch. We propose an extension of RDT in the sequential decision framework, namely Sequential RDT. Three classical signal deviation/distortion detection problems are reformulated and cast into the RDT framework. Using RDT and Sequential RDT, we investigate the detection of AutoPEEP (auto-Positive End Expiratory Pressure), a common ventilatory abnormality during mechanical ventilation. This is the first work of that kind in the state-of-the-art. Extension to the detection of other types of asynchrony is also studied and discussed. These early detectors of AutoPEEP and asynchrony are key elements of an automatic and continuous patient-ventilator interface monitoring framework.

Keywords : Hotspots, protein interaction, sequence-based features, frequency-based features, resonant recognition model, patient ventilator interaction, dynamic hyperinflation detection, AutoPEEP detection, random distortion testing, sequential decision.

Résumé

Cette étude présente des contributions en traitement statistique du signal avec des applications biomédicales. La thèse est divisée en deux parties.

La première partie traite de la détection des hotspots à l'interface des protéines. Les hotspots sont les résidus dont les contributions énergétiques sont les plus importantes dans l'interaction entre protéines. Les forêts aléatoires (Random Forests) sont utilisées pour la classification. Une nouvelle famille de descripteurs de hotspot est également introduite. Ces descripteurs sont basés seulement sur la séquence primaire unidimensionnelle d'acides aminés constituant la protéine. Aucune information sur la structure tridimensionnelle de la protéine ou le complexe n'est nécessaire. Ces descripteurs, capitalisant les caractéristiques fréquentielle des protéines, nous permettent de savoir la façon dont la séquence primaire d'une protéine peut déterminer sa structure tridimensionnelle et sa fonction.

Dans la deuxième partie, le RDT (Random Distortion Testing), un test robuste d'hypothèse, est considéré. Son application en détection du signal a montré que le RDT peut résister aux imperfections du modèle d'observation. Nous avons également proposé une extension séquentielle du RDT. Cette extension s'appelle le RDT Séquentiel. Trois problèmes classiques de détection d'écart/distorsion du signal sont reformulés et résolus dans le cadre du RDT. En utilisant le RDT et le RDT Séquentiel, nous étudions la détection d'AutoPEEP (auto-Positive End Expiratory Pressure), une anomalie fréquente en ventilation mécanique. C'est la première étude de ce type dans la littérature. L'extension à la détection d'autres types d'asynchronie est également étudiée et discutée. Ces détecteurs d'AutoPEEP et d'asynchronies sont les éléments principaux de la plateforme de suivi de manière automatique et continue l'interface patient-ventilateur en ventilation mécanique.

Mots-clés: Hotspots, interaction protéine, le modèle de résonance, l'interaction patient-ventilateur, détection d'AutoPEEP, random distortion testing, décision séquentielle.

General introduction

Motivation

In the recent years, biomedical engineering has been undergoing explosive growth. Many problems have been presented with challenges that require the cooperative efforts of scientists and engineers from different fields to provide a solution. Such multidisciplinary studies have become more and more important not only to conquer the challenge but also to enrich the competence of each domain. In Lab-STICC/CID/TOMS and the Signal and Communications department of TELECOM Bretagne, many techniques in information theory and statistical signal processing — in particular, sparsity, robust statistics, statistical detection/classification/estimation — have been developed. The purpose of this PhD research is to contribute to such expertise and know-how through applications in biomedical engineering. By confronting our in-house competence with that of experts of the application field, our endeavor is not only to solve the challenge in that domain but also to improve the theoretical background, the methodology and to transfer the knowledge to other possible application fields beyond the scope of biomedical engineering. In this work, statistical decision will be considered through two concrete applications: one, in biology, is the detection of protein interface hotspots and the other, in health care, is the monitoring of mechanical ventilation systems. The solutions to these two challenges with different information sources and requirements of the domains represents two non-parametric trends in statistical decision strategies.

Context and contributions

In signal processing, a statistical decision involves providing a choice among the possible hypotheses given the available observations, usually random variables or random vectors. On the basis of what information is supposed to be known on the observations, different approaches have been investigated in literature. Basically, they could be classified into two categories: parametric and non-parametric methods.

As long as the distributions of these random observations under each of the per-

missible hypotheses are known or, at least, can be estimated with a finite number of parameters, classical parametric statistical inferences can be used. With the available prior information on the conditional (under each hypothesis) distributions of the observations, these approaches provide the optimal solution to which the true hypothesis might actually be, given each observation. Neyman-Pearson and Bayes are currently the two main approaches for parametric hypothesis testing. Other approaches could be accounted, including: minimax, GLRT (generalized likelihood ratio test), invariance principle, etc. In the bayesian approach, *prior* probabilities of all the permissible hypotheses are supposed to be known. The decision is thus performed on the basis of the *posterior* probabilities, which are inferred through Bayes' law. No privilege is made on any of the hypotheses. The test is then optimized to minimize a global cost, generally the error probability. However, the choice of a *prior* distribution is usually subjective and therefore impractical in many applications. To avoid such requirement — also limitation — of the bayesian approach, the Neyman-Pearson approach considers the problem in terms of testing a baseline hypothesis, the so-called null hypothesis, against the alternative one. Based on the available conditional distributions of observations under the two hypotheses, the decision rule is formulated to maximize the detection probability while restricting the probability of false-alarm to some specified value. Such optimization is meaningful in many practical applications, usually the detection of rare events when the number of false-alarms must be under control. Intrusion detection by radar is among typical examples.

In practice, the distributions of observations are hardly known or estimated. However, either a dataset of observations under the hypotheses is available or partial information on observations is known prior to analysis. In such cases, non-parametric methods can be investigated. We consider two main trends in non-parametric statistics: machine learning and robust statistical hypothesis testing.

On the one hand, in the machine learning approach, to compensate the lack of prior knowledge on the conditional distributions of observations, it is required that a dataset is available and this dataset is sufficiently large to cover the characteristics of the observation distributions. A classifier will be trained on this dataset, called the *training set*, to optimize the decision based on what the classifier learns about how the observations (also called *features*, or *descriptors*) are organized in the feature space. Many classifiers have been proposed in the literature, including SVM (Support Vector Machine), RF (Random Forest), NNs (Neural Networks), k-NN (k Nearest Neighbors), etc. Since no statistical information on observations is required, descriptors can be extracted in as various ways as one can imagine. Machine learning is thus suitable for high dimensional data and problems in which our knowledge about the relationship between the considered hypotheses and the corresponding observations is still too limited to employ any proposition on the conditional distributions of the observations.

On the other hand, robust hypothesis testing, which requires as little information on the observation distribution as possible, is proposed. We consider a new approach in hypothesis testing, named Random Distortion Testing (RDT). Motivated by many practical situations encountered in signal processing, such as passive SONAR, cognitive radio, and Electromagnetic Compatibility (EMC), only the noise distribution is supposed to be known in this robust hypothesis testing framework. No knowledge on the clean signal is used. The test is optimized as in the Neyman-Pearson approach, i.e. to maximize the detection probability under a constraint on the false-alarm probability. The optimality criterion is based on invariance properties of the observation.

It could be noted that the parametric approaches bring statistical optimality whereas the non-parametric methods provide robustness. There are also methods, namely *semi-parametric* methods, that combine the advantages of both parametric and non-parametric ones. In some points of view, RDT could be considered as a semi-parametric method due to its robustness against variations of the signal distribution and its statistical optimality with respect to Neyman-Pearson's criteria.

This PhD research investigates the two aforementioned non-parametric statistical decision approaches through two problems in biomedical engineering.

In the first part, the detection of protein interface hotspots — the residues that contribute the most to the binding energy of a protein complex — is considered as an insightful application. This problem presents a challenge in biology. It is well-known that the protein structure decides its interaction with other molecules and thus, determines hotspots. However, how this relation is scripted is still far from our reach. It is then not possible to employ any assumption on the observations (protein structure) under each of the considered hypotheses (hotspot/non-hotspot). Fortunately, a database of protein hotspots/non-hotspots with energetic contribution measurements provided by experimental ASM (Alanine Scanning Mutagenesis) is available. Therefore, the machine learning approach can be used. Among various classifiers, Random Forest is considered for its high performance in terms of classification accuracy, its capability of dealing with unbalanced dataset, its possibility to be extended to unlabeled data and, especially, its robustness in terms of tuning parameters. The main contribution of this part is twofold.

- First, the success in identifying protein hotspots using RF has provided a better insight into the nature of protein hotspots and protein interaction.

- Second, new descriptors based solely on protein primary amino-acid sequence are proposed. These descriptors are as discriminant as those derived from the 3-dimensional structure of the protein and/or the complex. They have also been shown to improve the hotspot detection performance when being combined with those from the 3-dimensional structure.

In the second part, the RDT robust hypothesis testing is investigated throughout the development of a universal framework for respiratory system monitoring. More specifically, the detection of AutoPEEP (auto Positive End-Expiratory Presure) and asynchrony during mechanical ventilation based on respiratory signals is addressed. These signals include respiration flow, air volume, airway pressure, etc available on most of the currently used ventilators. By this study, our purpose is to develop a system capable of incorporating expertise of the therapist into the detection process and to provide him a functionality to optimize his care. The approach proposed in this work is very general and could be used in many other application fields, including fault detection and structural health monitoring. Our main contributions are threefold.

- First, to the best of our knowledge, this is the first work on automatic detection of AutoPEEP, one of the most frequent abnormalities at the patient-ventilator interface during mechanical ventilation. The proposed detection of asynchrony is one of the two studies that can be found in the state-of-the-art literature.
- Second, three signal deviation/distortion detection problems have been identified and solved using RDT. These problems are very general and can be encountered in many application fields.
- Third, the RDT has been extended to sequential analysis, which improves the detection performance whilst still respecting the expectation for the probability of false-alarm.

This work has won the supports from Institut Mine-Telecom and Region Bretagne to study business aspects for a Spin-Off.

Organisation of the thesis

With respect to the content of this PhD research, the thesis will be organized in two parts:

- In Part 1, the detection of protein interface hotspot will be investigated. This part begins with a short summary of the non-parametric statistical decision in the Machine Learning framework. As one of the most powerful state-of-the-art classifiers, Random Forest is briefly introduced. The protein hotspot detection is then addressed with new descriptors proposed. Finally, the results and some further discussions will conclude this part.
- In Part 2, a universal framework for respiratory system monitoring is proposed. To begin with, RDT is briefly summarized with main results. Three deviation/distortion detection problems are also introduced, including: detection of

deviation at critical instants, change point detection and detection of distortion with respect to some reference during an interval of time. These three generalized problems are then addressed and tackled using RDT. The extension of RDT in sequential analysis framework is also presented. As a motivating application, the detection of AutoPEEP and asynchrony during mechanical ventilation is then investigated before discussing implementation in the monitoring framework. Experimental results and clinical analysis are also reported. The part is concluded with a discussion and some perspectives.

The thesis ends up with general conclusions on this PhD research. A discussion on several practical aspects of parametric and non-parametric statistical decision will be included. Some further perspectives on trends that this research might initiate will also be discussed.

List of publications:

- **Journals:**

  1. Q.-T. Nguyen, R. Fablet, and D. Pastor,"Protein interaction hotspot identification using sequence-based frequency-derived features", *IEEE Transactions on Biomedical Engineering*, Issue 99, July 2011.
  2. Q.-T. Nguyen, D. Pastor, and E. L'Her, "Automatic detection of AutoPEEP during controlled mechanical ventilation", *BMC Biomedical Engineering On-Line*, vol. 11, no. 1, p. 32, 2012. [Highly accessed]
  3. D. Pastor, and Q.-T. Nguyen, "Random distortion testing and optimality of thresholding tests", *IEEE Transactions on Signal Processing*, 2012. [Submitted]

- **Conferences:**

  1. Q.-T. Nguyen, D. Pastor, and E. L'Her, "Patient-ventilator asynchrony: automatic detection of AutoPEEP", ICCASSP 2012: 37th International Conference on Acoustics, Speech, and Signal Processing , 25-30 march, Kyoto, Japan, 2012.
  2. E. L'Her, Q.-T. Nguyen, D. Pastor, and F. Lellouche, "Automatic flow curves analysis during mechanical ventilation (CURVEX): Application to Intrinsic PEEP detection", ATS 2012 : International Conference of the American Thoracic Society, 18-23 May, San Francisco, California, USA, 2012.

- **Patent:**

1. Q.-T. Nguyen, D. Pastor, E. L'Her, and F. Lellouche, "Procédé de détection d'au moins une anomalie dans un signal observé, produit programme d'ordinateur et dispositif correspondants", brevet français, No. 12/52660, déposée le 23 mars 2012.

- **Technical reports:**

1. Q.-T. Nguyen, D. Pastor, and E. L'Her, "A statistical signal processing platform for automatic detection of pulmonary distension during assisted mechanical ventilation", Research Report, Institut Télécom, Télécom Bretagne, RR-2012-02-SC, 2012.
2. D. Pastor and Q.-T. Nguyen, "Testing the Mahalanobis distance between a random signal with unknown distribution and a known deterministic model in additive and independent Gaussian noise: the random distortion testing problem" (3rd version), Research Report, Institut Télécom, Télécom Bretagne, RR-2012-04-SC, 2012.

Part I

Statistical decision in Machine Learning framework Application: Protein interface hotspot detection

CHAPTER 1

Statistical decision in Machine Learning - Random Forests (RF)

As aforementioned, prior knowledge on the conditional distributions of observations under each of the permissible hypotheses is hardly known in practice. However, in many real-world applications, gathering data is a much simpler task. Therefore, the question on how the relationship between observations and its corresponding hypothesis can be learned from available dataset arises. The machine learning framework is introduced to address such question. Given an available dataset, called *training set*, $\mathcal{L} = \{(\mathbf{v}^i, h^i), i = 1..N\}$ of observations $\mathbf{v}^i \in \mathcal{V}$ and the corresponding hypotheses $h^i \in \mathcal{H}$ where $\mathcal{V}$ is the feature space and $\mathcal{H}$ is the hypothesis space, the machine learning is aimed at estimating — on the basis of $\mathcal{L}$ — a function $\varphi : \mathcal{V} \to \mathcal{H}$ that maps any element of feature space $\mathcal{V}$ to a hypothesis of $\mathcal{H}$. Depending on what is known on the true hypotheses $h^i, i = 1..N$ of the data instances $\mathbf{v}^i, i = 1..N$ in the training set $\mathcal{L}$, the machine learning could be divided into: unsupervised [Edward and Cavalli-Sforza, 1965, Hartigan, 1975, Dempster et al., 1977], supervised [Dempster et al., 1977, Vapnik, 1995, Freund and Schapire, 1997, Breiman, 2001], semi-supervised [Chapelle et al., 2006], presence/absence [Shivani and Roth, 2002, Opelt et al., 2004, Ulusoy and Bishop, 2005], weakly-supervised learning [Lefort, 2010], etc. And depending on the considered hypotheses, a machine learning scheme can serve different tasks, including clustering, classification and regression (see [Bishop, 2006] amongst others). Since our focus is on statistical decision, the supervised learning-based classification is of great interest. In other words, we consider the case where $h^i, i = 1..N$ are supposed to be known for all data instances $\mathbf{v}^i, i = 1..N$ of the training set and $\mathcal{H}$ is a finite set of permissible hypotheses. Among the most powerful state-of-the-art classifiers, one could mention SVM (Support Vector Machine) [Cortes and Vapnik, 1995], RF (Random Forests) [Breiman, 2001], and NNs (Neural Networks) (see [Bishop, 1996, Bishop, 2006] amongst others). These classifiers have been successfully used in many real-world applications, including: speech recognition, computer vision, bio-surveillance, remote sensing, robot-control, bio-informatics, etc.

In this work, the Random Forest is used for its high classification performance, its robustness to variation and its simplicity in terms of empirical parameters tuning. This chapter begins with a brief introduction to classification tree, the elementary component of any random forest. The bagging predictor, the technique that makes it possible to combine multiples weak predictors to form a powerful one, is then presented. Finally, the random forest is summarized.

1.1 Classification tree

A classification or decision tree [Breiman et al., 1984] is a tree-structured predictive model in which each internal node is associated with a decision rule based on object features $\mathbf{v} = [v_1, v_2, ..., v_D]^T \in \mathcal{V}$ and each terminal leaf is assigned to a class $h \in \mathcal{H}$ ($\mathcal{H} = \{0, 1\}$ for a binary classification problem such as hotspot identification). Given a decision tree, the class of an object is predicted by filtering its features through the successive decision rules of the internal nodes until a terminal leaf is reached. The class of the terminal leaf is then assigned to the object. In the considered random forest setting, decision trees are binary and the decision rule at each internal node of the tree is a test on only one of the object features, say j-th feature. In this test, v_j is compared to its associated threshold λ_j. Objects with v_j less than the threshold λ_j will be filtered to the left child node and the others — with v_j greater than λ_j — are forwarded to the right one. Figure 1.1 shows an example of a decision tree.

The construction of a binary decision tree is generally performed on the basis of a training set $\mathcal{L}$. Starting from the root node with all the training samples $\{(\mathbf{v}^i, h^i), i = 1..N\}$, the decision tree is grown by recursively splitting nodes in such a way that at each node t_p, the training samples are divided into two subsets (corresponding to two children nodes, t_L and t_R) with maximum class homogeneity according to a decision rule. The determination of the decision rule associated with each split amounts to seeking the best feature and its best threshold that maximize the information gain G:

$$(j, \lambda_j)^* = \arg\max_{j,\lambda_j} G$$

where the information gain G yielded by the division of samples in a parent node t_p into two child nodes t_L and t_R is defined by:

$$G = I(t_p) - p_L I(t_L) - p_R I(t_R),$$

in which p_L (resp. p_R) is the fraction of samples in t_p that will be sent to the child left node t_L (resp. the child right node t_R) and $I(t)$ is the *impurity* of node t [Breiman et al., 1984]. For binary classification problems, node impurity can be interpreted as the proportion of the less frequent class in the sample subset associated with that node.

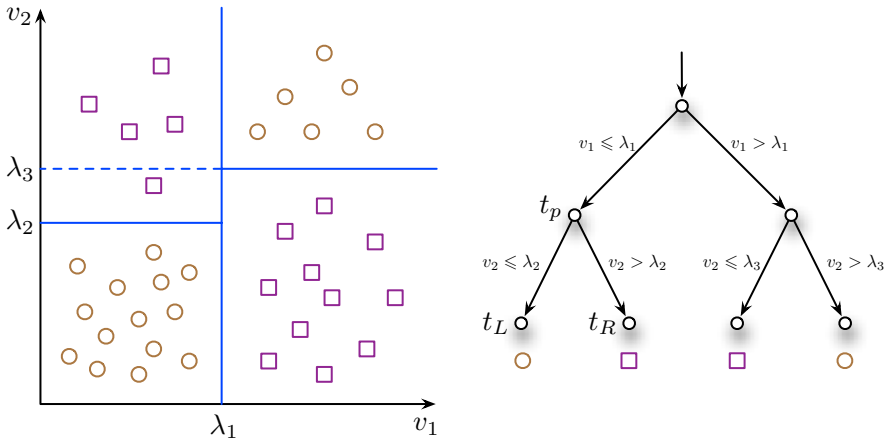


Figure 1.1 — Example of a classification tree. This example involves a two-class dataset of objects in a 2-dimensional feature space (left). From training samples of each class, represented as squares and circles in the left figure, the classification (decision) tree in the right figure is built. The solid lines in the left figure show the division of the feature space into homogeneous regions in which only samples of the same class are present.

The perfect split is therefore the one that divides all samples of the parent node into two separate classes. In practice, to measure node impurity, Shannon's entropy

$$I(t) = -\sum_k p_k^{(t)} \log_2(p_k^{(t)})$$

and Gini's diversity index

$$I(t) = \sum_k p_k^{(t)}(1 - p_k^{(t)}),$$

where $p_k^{(t)}$ is the proportion of samples of class k (i.e. hypothesis h_k) in node t, are usually used [Breiman et al., 1984, Quinlan, 1993]. Using the aforementioned splitting rules, the decision tree is recursively grown until maximum homogeneity, i.e. minimum impurity, is obtained in the terminal leaf nodes. The construction of a decision tree can be regarded as an adapted quantization of the feature space $\mathcal{V}$ into homogeneous regions, in which most training samples are of the same class — and this class will be assigned to any new sample observed in that region (cf. Figure 1.1).

1.2 Bagging predictors

Bootstrap **agg**regat**ing** (bagging) is a machine learning ensemble technique to improve the prediction performance by generating multiple versions of a considered predictor and using these to form an aggregated one, which can reduce the variance and also avoid over-fitting [Breiman, 1996]. This technique was first proposed by Leo Breiman

with successful demonstrations on classification and regression trees. The bagging technique can be resumed in this section as follows.

Let $\mathcal{L} = \{(\mathbf{v}^i, h^i), i = 1..N\}$ be the training set, where h's are the true hypotheses — or the class labels in classification — of the data instances and $\varphi_{\mathcal{L}} : \mathcal{V} \to \mathcal{H}$ be the predictor (classifier) learned from $\mathcal{L}$ by some procedure, for example the aforementioned classification tree. Now supposing that a sequence of training sets $\{\mathcal{L}_k\}$, each consisting of N data instances drawn from the same distribution as that of $\mathcal{L}$, are available for training a sequence of predictors $\{\varphi_{\mathcal{L}_k}\}$. It is then questioned that how $\{\varphi_{\mathcal{L}_k}\}$ can be used to get a more powerful predictor than the single training set one $\varphi_{\mathcal{L}}$. The aggregating technique suggested in Breiman's bagging addresses such question. It has been shown in [Breiman, 1996] that by voting, the aggregated classifier φ_A defined by:

$$\varphi_A(\mathbf{v}) = \arg\max_j \ \#\{k : \varphi_{\mathcal{L}_k}(\mathbf{v}) = j\}, \tag{1.1}$$

where # represents an operator that counts the number of instances in which its argument is true, yields better classification performance than $\varphi_{\mathcal{L}}$.

However, in practice, only one training set $\mathcal{L}$ is available. It is then required that $\{\mathcal{L}_k\}$ can be generated from $\mathcal{L}$. The Breiman's bootstrap performs such task. Given $\mathcal{L}$, each replicate dataset $\mathcal{L}_k^{(B)}$ can be constructed by randomly drawing, *with replacement*, N data instances from $\mathcal{L}$. By such a process, each $(\mathbf{v}^i, h^i)$ might appear repeated times or not at all in any bootstrap dataset $\mathcal{L}_k^{(B)}$. The sequence $\{\mathcal{L}_k^{(B)}\}$ of replicate bootstrap datasets, each consisting of N sample drawing from the bootstrap distribution approximating the distribution underlying $\mathcal{L}$ [Breiman, 1996], can then be used in place of $\{\mathcal{L}_k\}$ for training individual predictors $\{\varphi_{\mathcal{L}_k^{(B)}}\}$. The resulting aggregated classifier is then given by:

$$\varphi_A^{(B)}(\mathbf{v}) = \arg\max_j \ \#\{k : \varphi_{\mathcal{L}_k^{(B)}}(\mathbf{v}) = j\} \tag{1.2}$$

1.3 Random Forests

Based on bagging technique, Random Forest [Breiman, 2001] is an ensemble classifier that combines *nbTrees* decision trees. These trees are constructed using subsets $\{\mathcal{L}_k^{(B)}\}$ of individuals that are independently and randomly sampled from the original training set $\mathcal{L}$. The search for the optimal splitting rule of each node is optimized with respect to a randomly selected subset of features. The classification of an input is obtained by aggregating the votes of the individual trees in the forest. By combining two sources of randomness, i.e. the random selection of training samples and the random selection of features for the determination of each splitting criterion, classification performance of RF greatly improves compared to a single decision tree [Breiman, 2001].

As aforementioned, a decision tree can be considered as a partition of the feature

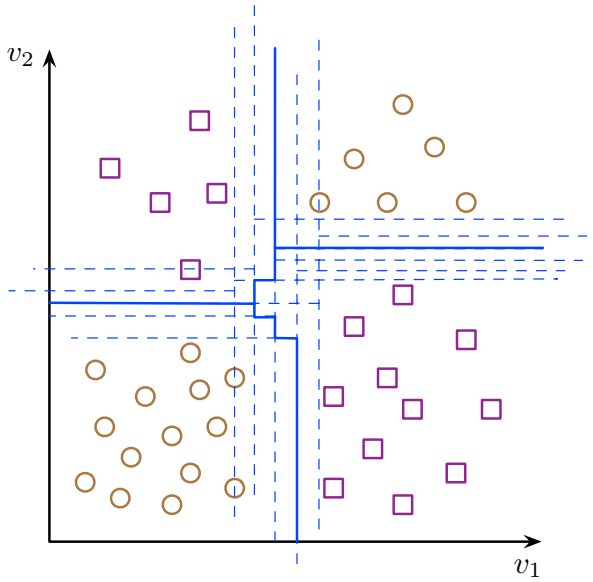


Figure 1.2 — The optimal partition of feature space $\mathcal{V}$ into homogeneous regions provided by Random Forest. In this example, the binary classification with two descriptors v_1 and v_2 is considered. The dash lines represents the intermediate variated partitions yielded by individual decision trees. The solid lines point out the optimal partition resulted from RF by aggregating those from individual trees.

space $\mathcal{V}$ into homogeneous regions on the basis of training set. By using replicate bootstrap training sets $\mathcal{L}_k^{(B)}$, each of the constructed trees presents a variated version of partition resulted from $\mathcal{L}$. This variation, driven by the randomness of the bootstrap samples, plays a very important role in avoiding over-fitting. By aggregating with majority vote, the resulting RF thus presents a better partition of the feature space $\mathcal{V}$ than any individual decision trees, including the one yielded by the original training set $\mathcal{L}$. Figure 1.2 illustrates that fact.

RF has been shown to be among the most efficient machine learning schemes for a variety of issues, such as mass spectrometry data analysis [Wu et al., 2003], microarray data analysis [Diaz-Uriarte and Alvarez de Andres, 2006], protein interaction prediction [Chen and Liu, 2005], network security [Zhang et al., 2008], language modeling [Xu and Jelinek, 2007], image recognition [Nguyen et al., 2012], etc. The construction of an RF involves only two parameters: the number *nbTrees* of trees and the number *mTry* of randomly selected features for the determination of each optimal splitting criterion. These key characteristics make RF a good choice for hotspot data and, particularly, for our purpose of assessing and comparing the relevance of the descriptors.

CHAPTER 2

Protein interface hotspot detection

2.1 Introduction

Understanding the structure and the biological function of proteins, the elementary building blocks of all living organisms, is among the most important topics in biology [Alberts et al., 2010]. Basically, biochemists refer protein structure in four distinct levels. In primary structure, a protein is constituted of amino acids organized in a linear sequence. The content and the order of this sequence is determined by the gene corresponding to the protein. In the second level, when hydrogen bonds are established among the main-chain peptide groups, the amino acid sequence is folded to form regular local sub-structures — such as α-*helices*, β-sheets — called secondary structure. Driven by hydrophobic interactions, in the third level, α-*helices* and β-sheets are folded into a compact globule. This three-dimensional structure of a single protein molecule is referred to as the tertiary structure. For proteins consisting of more than one amino acid chain, there exists a quaternary structure, which describes how the subunits interact and fit together in a three-dimensional form. In the recent years, scientists are working together to answer the question on how the primary amino acid sequence of the protein defines its conformation and function [Alberts et al., 2010, Cosic, 1997, Ofran and Rost, 2007]. Solving this issue could open a new era in biology where most bioactivities can be controlled, including curing diseases by newly designed proteins with pre-defined functions (see [Cosic, 1997] amongst others).

Studies in biology have shown that proteins form certain active three-dimensional structures to interact with other molecules through their interfaces [Alberts et al., 2010]. Actually, "most interfaces are composed of two relatively large protein surfaces with good shape and electrostatic complementarity for one another" [Bogan and Thorn, 1998]. It has also been shown that the distribution of binding energies on these interfaces is not uniform [Bogan and Thorn, 1998]. Some residues are more important than others as they comprise only a small fraction of the interface but contribute most of the necessary energies to the interaction [Ofran and Rost, 2007]. If they are mutated, the

interaction may be affected and, as a result, the protein function may be altered. These critical residues are commonly referred to as *hotspots* [Wells, 1991, Bogan and Thorn, 1998]. Figure 2.1 shows an example of such protein hotspots. The characterization, detection and identification of *hotspots* are then keys to the understanding of protein interactions and functions. Much research, both experimental and computational, has been conducted to shed light on these critical residues of the interfaces [Bogan and Thorn, 1998, Kortemme et al., 2004, Guerois et al., 2002, Gao et al., 2004, Rajamani et al., 2004, Brinda et al., 2002, Darnell et al., 2007, Ofran and Rost, 2007, Cho et al., 2009, Fernández-Recio et al., 2004, Ramachandran et al., 2004, Ramachandran and Antoniou, 2008, Deergha Rao and Swamy, 2008, Sahu and Panda, 2009].

Experimentally, the energy contribution of a given residue to the interaction of a protein with its target can be determined by measuring the change in binding free energy when this residue is *in vitro* mutated to alanine. When the measured change in binding free energy is large enough, this residue is deemed as a hotspot [Wells, 1991]. This method, also known as alanine scanning mutagenesis (ASM), was used by Thorn and Borgan to analyze hotspots and the database that they established is referred to as the Alanine Scanning Energetics database (ASEdb) [Thorn and Bogan, 2001]. Unfortunately, such a widely accepted experimental method requires significant effort and hence induces low throughput [Ofran and Rost, 2007, Kortemme et al., 2004].

In the search for lower-cost methods applicable to high-throughput analysis, computational approaches have been proposed to identify hotspot residues in protein interfaces. In this respect, Kortemme and Baker [Kortemme and Baker, 2002] introduced a simple physical model for binding free energy. This model takes into account packing interaction, polar interaction involving ion pairs and hydrogen bonds, and solvation. Hotspots are then identified by computational alanine scanning (Robetta) [Kortemme et al., 2004], which involves the numerical evaluation of the change in this binding free energy of protein-protein complexes due to computational alanine mutations. These computationally identified hotspots are shown to be in agreement with those identified by *in vitro* experiments and reported in the ASEdb database. Motivated by these works, other energy-based methods have been proposed in [Guerois et al., 2002, Gao et al., 2004]. Other computational approaches also investigated molecular dynamics (MD) simulations [Rajamani et al., 2004], graph analysis [Brinda et al., 2002] and machine learning [Ofran and Rost, 2007, Darnell et al., 2007, Cho et al., 2009]. Among all the aforementioned methods, the most successful ones require the structure of the complex — or, at least, the three-dimensional structure of the protein — to be known. The docking approach in [Fernández-Recio et al., 2004], which requires simulating thousands of possible docking poses for the protein complex, is among the most popular in this respect.

Although the biological functions of proteins relate to certain active tertiary struc-

tures, it is assumed that all information about their structures and, thus their functions, is primarily embedded in amino acid sequences [Alberts et al., 2010]. In other words, knowledge of the three-dimensional structure of the protein or of the complex is expected to be more than sufficient to identify hotspots of the interfaces. In [Ofran and Rost, 2007], Ofran and Rost showed that hotspots can probably be predicted using only amino acid sequence information. Albeit less accurate than methods based on available three-dimensional (3D) structure information, their sequence-based hotspot identification method yielded relevant results. On the other hand, the introduction of the Resonant Recognition Model (RRM) by I. Cosic in [Cosic, 1994] pointed out the existence of a *characteristic frequency*, which represents a certain periodicity within the energy distribution of valence electrons along the protein molecule. This finding has inspired many attempts to detect hotspots by using digital signal processing (DSP) methods, such as those based on Short-time Fourier transform (STFT) [Ramachandran et al., 2004], digital filters [Ramachandran and Antoniou, 2008], wavelet transform [Deergha Rao and Swamy, 2008] and S-Transform filtering [Sahu and Panda, 2009]. Though tested on only a few individual sequences, these approaches suggest time-series analysis as a relevant framework for hotspot identification.

In this study, we suggest a learning-based hotspot identification using Random Forest. A new family of frequency-based descriptors derived solely from the protein primary amino acid sequence is also proposed. These descriptors are extracted using a simple *in silico* alanine scanning and DSP techniques based on the discrete Fourier transform. The relevance of the proposed descriptors is evaluated through a machine-learning-based classification. The underlying idea is that once a classifier successfully separates hotspots from non-hotspots via certain given features, these features are then considered to be capable of discriminating hotspots from non-hotspots. In other words, these features are actually relevant to the hotspot identification problem. In this respect, as one of the most powerful machine learning algorithms [Breiman, 2001], Random Forest is used for its advantages with respect to state-of-the-art, in particular, its high classification performance, its robustness to variation and its simplicity in terms of empirical parameter tuning. Other approaches, such as SVM (Support Vector Machine), could also be considered; however, it should be kept in mind that the detection results might be very sensitive to the choice of parameters.

The results on the dataset show that the proposed frequency-based descriptors can be used to achieve an accuracy of 79% and a precision of 75%. Without information on the three-dimensional structure of the protein and/or the complex, these descriptors can achieve performance comparable to that reported in [Kortemme et al., 2004] and [Tuncbag et al., 2009] where such information is required. This is a key feature since knowing the protein 3D structure, either computationally or experimentally, is not straightforward, and actually, most protein sequences are available without 3D

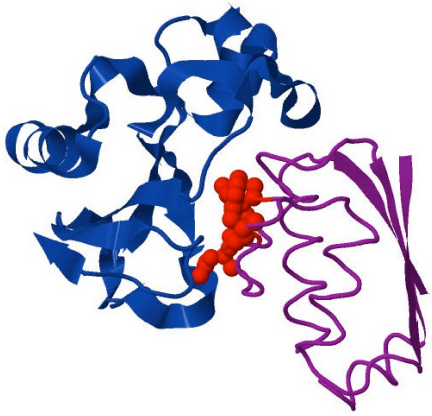


Figure 2.1 — An example of a protein with hotspots. In this figure, the barstar molecule (right/violet) with hotspots (red) is shown to be in interaction with barnase (left/blue), forming the complex barstar-barnase. The three-dimensional structures of barstar and its target, barnase, are represented in terms of basic secondary structure motifs (α-helices, β-sheets, turns) while red balls indicate atoms of hotspot residues. The structure of the complex was retrieved from the Protein Data Bank (PDB) using its identity 1brs. On the other hand, information on the hotspot residues involved in this interaction was provided by ASEdb.

structure information. The experimental results also show that our sequence-based frequency-derived descriptors can boost the prediction up to an accuracy of 82% and a precision of 80% when combined with the 3D structure-based features proposed in [Tuncbag et al., 2009]. Moreover, using DSP techniques on one-dimensional sequences, our method requires very little computational load and thus can be applied to large-scale analysis.

2.2 Sequence-based frequency-derived features

2.2.1 Conversion to numerical sequence

The primary structure of a protein is given by the associated sequence of amino acids. This sequence is often represented by a string of characters sampled from an alphabet of 20 single characters representing the 20 different amino acids. By properly mapping these character strings into numerical sequences, time series analysis can be applied to design very high throughput methods. This conversion from symbolic to numerical sequences may rely on assigning to each amino acid numerical values that represent its physico-chemical and biochemical properties. A number of such indices have been introduced in the literature (more than 500 indices can be found in the AAIndex database [Kawashima et al., 2008]). Among them, the *electron-ion interaction pseudo-potential* (EIIP) values [Cosic, 1994] and the *ionization constant* (IC)

parameters [Cosic and Pirogova, 1998] are shown to be very relevant to the protein bioactivity. For each amino acid, the EIIP value describes the average energy states of all valence electrons of its atoms. This can be calculated using the general model of pseudo-potential [Veljkovic and Slavic, 1972]:

$$< \overrightarrow{k+q}|w|\overrightarrow{k} >= 0.25\bar{Z}\sin(\pi 1.04\bar{Z})/(2\pi) \tag{2.1}$$

where q is a change of momentum k of the delocalized electron in the interaction with potential w and $\bar{Z}$ is the average number of valence electrons of an atom. Let us take the calculation of the EIIP value for Asparagine (ASN) for example. Its residue (-CH_2CONH_2) is composed of 2 carbon (C), 1 oxygen (O), 1 nitrogen (N) and 4 hydrogen (H) atoms. Therefore, the average number of valence electrons per atom is $\bar{Z} = (2\times 4 + 1\times 6 + 1\times 5 + 4\times 1)/(2+1+1+4) = 23/8$. By substituting this value into the formula (2.1) to compute the pseudo-potential, the EIIP value for Asparagine (ASN) is then 0.0036. The IC value of an amino acid $H-A$ measures its acid dissociation constant from the corresponding ionization reaction $H-A = H^+ + A^-$, computed as follows:

$$\text{pKa} = -\log_{10}\text{Ka} \tag{2.2}$$

with

$$\text{Ka} = \frac{[H^+][A^-]}{[H-A]} \tag{2.3}$$

where $[H^+]$, $[A^-]$ and $[H-A]$ are respectively the concentration of positively charged ions, negatively charged ions and reactant in the solution. The EIIP and IC values for the 20 amino acids occurring in nature are listed in Table 2.1. These two indices have been shown to be very successful in the so-called Resonant Recognition Model [Cosic, 1994, Cosic, 1997, Cosic and Pirogova, 1998] (cf. Section 2.4.2) to get an insight into the physical characterization of protein interactions as well as protein hotspots. In our work, these indices will be used to obtain numerical sequences for further DSP analysis.

2.2.2 In-sillico alanine scanning and frequency-based features

Experimental alanine scanning mutagenesis has been shown to be an extremely useful tool for analyzing interactions in protein interfaces (see [Wells, 1991, Kortemme et al., 2004] amongst others). This technique involves mutating an amino acid residue to alanine (i.e. deleting the sidechain beyond C_β carbon atom) and then evaluating the effects of this mutation on the affinity of the protein interaction. These effects can be measured by the change in binding free energy ($\Delta\Delta G$) of the protein-target complex. Although experimental ASM is very powerful in identifying hotspot residue, it is still too expensive and laborious to be easily applied to large-scale analysis, despite many advances in molecular biology.

Table 2.1 — EIIP and IC numerical values

Amino acid	3-Letter Code	1-Letter Code	EIIP (residue)	IC
Leucine	LEU	L	0.0000	2.4000
Isoleucine	ILE	I	0.0000	2.4000
Asparagine	ASN	N	0.0036	2.2000
Glycine	GLY	G	0.0050	2.4600
Valine	VAL	V	0.0057	2.3500
Glutamic Acid	GLU	E	0.0058	2.3000
Proline	PRO	P	0.0198	2.0000
Histidine	HIS	H	0.0242	2.3000
Lysine	LYS	K	0.0371	2.2000
Alanine	ALA	A	0.0373	2.3000
Tyrosine	TYR	Y	0.0516	2.2000
Tryptophan	TRP	W	0.0548	2.3700
Glutamine	GLN	Q	0.0761	2.0600
Methionine	MET	M	0.0823	2.1700
Serine	SER	S	0.0829	2.1000
Cysteine	CYS	C	0.0829	1.9600
Threonine	THR	T	0.0941	2.0900
Phenylalanine	PHE	F	0.0946	1.9800
Arginine	ARG	R	0.0959	1.8200
Aspartic Acid	ASP	D	0.1263	1.8800

Here, we investigate an alternative based on a purely computational approach. More specifically, we propose an *in silico* alanine scanning approach inspired from the experimental ASM. We proceed as in ASM, but computationally, by replacing subsequences of residues by alanines and looking for frequency-related changes in the overall sequence. The approach is very similar to the computational alanine scanning method described in [Kortemme et al., 2004]. However, instead of investigating a physical model or a single measure that relates to binding free energy as in [Kortemme et al., 2004], we analyze changes in the frequency spectrum caused by computational mutagenesis.

The proposed framework is sketched in Figure 2.2. Our alanine scanning module computationally mutates residues around a given position j of the input amino acid sequence $s(n)$ to alanines. Instead of replacing residue $s(j)$ only, a window of residues centered at position j is processed. All the residues of the window are thus computationally mutated to alanines since changing the value of one single sample will not significantly affect the spectrum of the sequence. On the other hand, the O-ring theory also claims that hotspots are surrounded by other residues, less important in binding energy, but whose role is likely to occlude bulk solvent from central residues to form high affinity interactions [Bogan and Thorn, 1998]. To take these surrounding residues into account, a window of length $L = 5$ — the tested residue $s(j)$ itself and two residues on each side — has been empirically chosen. Furthermore, this choice is reasonable with

respect to cases where hotspots are very close to each other.

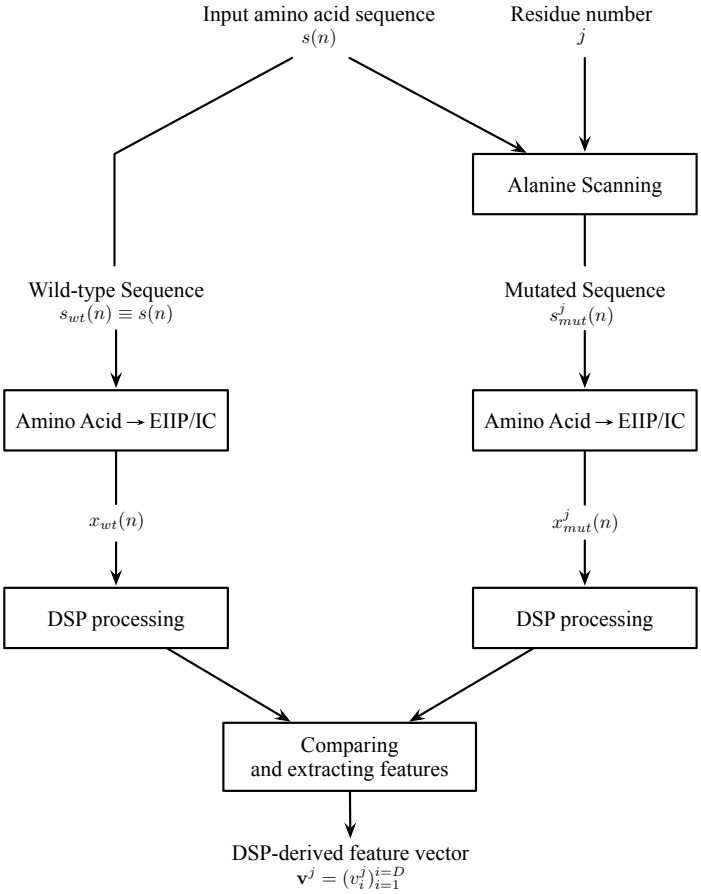


Figure 2.2 — Computational alanine scanning and DSP-based features deriving

After computational mutation, both the wild-type sequence $s_{wt}(n)$ and the mutated one $s^j_{mut}(n)$ are converted into numerical sequences ($x_{wt}(n)$ and $x^j_{mut}(n)$, respectively) using either EIIP or IC values. These two numerical sequences will then be analyzed by the same DSP scheme and their associated frequency-based characteristics will be further compared to derive the proposed descriptor vector $\mathbf{v}^j$. Various DSP techniques, both traditional and modern, may be considered, including Fast Fourier Transform (FFT), Short-time Fourier transform (STFT) or wavelet transform. Similarly, many characteristics could be considered, including peak frequencies, sub-bands energies, and so on. Within our framework, as comparison criteria, we focus on spectrum peak changes, sub-band energy changes and global energy changes. These features can be regarded as the analysis at different levels of resolution, from local to global, of the frequency spectrum.

Spectrum peak changes

Both the wild-type and the computationally mutated numerical sequences (i.e. $x_{wt}(n)$ and $x^j_{mut}(n)$, respectively) are transformed into the frequency domain by FFT. Peak frequencies are defined as the local maximum points of the wild-type sequence frequency amplitude spectrum. For discrete sequences, we define the set I of these peak frequencies as:

$$I = \{0 < k < N : |X_{wt}(k)| \geqslant \max(|X_{wt}(k-1)|, |X_{wt}(k+1)|)\}$$

where $X_{wt} = FFT(x_{wt})$ is the FFT of x_{wt} and the FFT size N is chosen to be equal to the sequence length. The DC component is removed from the input sequence before FFT to avoid any spurious peak at the null frequency. Since the amplitude spectrum is symmetric, only one half of it is considered. In terms of the RRM [Cosic, 1997], these peak frequencies are regarded as potential characteristic frequencies of the protein functions (cf. our discussion in Section 2.4.2). Changes in the amplitude spectrum at peak frequencies caused by computational mutation are regarded as potential signatures of hotspots. More precisely, we compute the following features:

$$PeakChange^j_k = \frac{|X_{wt}(k)|}{|X^j_{mut}(k)|}$$

where $X^j_{mut} = FFT(x^j_{mut})$ and k is among the considered peak frequencies. In this study, only the set of the three highest peak changes will be retained and will be taken as descriptors.

Sub-band energy changes

In addition to amplitude changes at peak frequencies, local energy-changes in frequency subbands are also considered. Specifically, sequences are transformed into time-frequency representations using STFT with a sliding window of length $(\frac{N}{4} + 1)$, where the number N of FFT points is now chosen to be the smallest power of two greater than or equal to the sequence length. This value is the default configuration of the Time-Frequency Toolbox (`http://tftb.nongnu.org/`) that we use to perform time-frequency analysis. To achieve a relevant time-frequency analysis, an analyzing window with small side-lobes is required. According to [Harris, 1978], the 4-term Blackman-Harris window is adopted here for its trade-off between the main-lobe width and the side-lobe levels. Other windows with low side-lobe levels such as the Blackman and the Gaussian windows were also tested and provided similar results. Moderate windows, such as Hamming and Hanning, were shown to be less efficient. After the STFT, since the frequency spectra $S^j_{mut}(j, .)$ and $S_{wt}(j, .)$ at mutated position j are also symmetric, the higher halves can be discarded. The retained lower halves are then evenly divided

into 8 equal sub-bands. The change in energy due to the computational mutation will be considered in these 8 sub-bands by computing

$$SBEnergyChange_m^j = \frac{\sum_{\nu \in SB_m} |S_{wt}(j,\nu)|^2}{\sum_{\nu \in SB_m} |S_{mut}^j(j,\nu)|^2}, m = 1..8$$

where

$$S_{wt} = STFT(x_{wt})$$
$$S_{mut}^j = STFT(x_{mut}^j)$$

and SB_m is the m-th sub-band

$$SB_m = \{k : (m-1)\frac{N}{16} \leqslant k < m\frac{N}{16}\}.$$

Global energy changes

Global energy change is defined as the ratio of the mutated sequence energy to that of the wild-type one:

$$EnergyChange^j = \frac{\sum_{n=1}^{L} |x_{mut}^j(n)|^2}{\sum_{n=1}^{L} |x_{wt}(n)|^2}$$

where L is the sequence length. Of course, this energy ratio can be equivalently computed in the frequency domain.

2.3 Learning-based hotspot identification

To computationally detect hospot, the learning-based recognition scheme is suggested. In this study, we exploit Random Forest (RF) [Breiman, 2001] as the learning-based classifier since it is among the most powerful techniques for supervised classification issues. The detection is carried out on the basic of two different families of protein hotspot descriptors: the proposed features derived from frequency characteristics of the protein's amino acid sequence and state-of-the-art features computed from known 3D structure of the considered proteins and/or the complexes. These two families of descriptors can be used separately or together upon the availability of the prerequisite knowledge on the 3D structure. The evaluation is also carried out on a dataset with a comparison of detection performance yielded by each of the feature families and the combination to illustrate, on the one hand, the relevance of the proposed descriptors for hotspot identification, and, on the other hand, the success of the proposed detection in the machine learning framework. To begin with, the evaluated features are pointed out. The considered hotspot dataset is then presented. And finally, the detection results are reported.

2.3.1 Evaluated features

As aforementioned, we consider the two following sets of protein hotspot descriptors as the input of the RF classifier in the learning-based hotspot identification scheme.

Frequency-derived features of amino acid sequences

The frequency-based features presented in Section 2.2.2, that is, the 3 highest spectrum peak changes, the 8 sub-band energy changes and the global energy changes, are considered. Using these measures with both EIIP and IC values, a set of 24 different features is computed. The descriptors that best discriminate hotspots from other residues will be selected. This can help reduce the dimensionality of the feature space, without affecting the original semantics of the descriptors, thus providing the ability to interpret the result by domain experts [Saeys et al., 2007]. In this study, such a selection is performed by using a decision tree-based feature ranking technique [Cardie, 1993]. The technique involves growing a decision tree based on a sample set (cf. section 1.3 for more details) then pruning it at a certain level. During the growing process, a decision tree, by its nature, selects the best feature (in the sense of maximizing the information gain) each time a node is split. In the pruning phase, nodes that provide less entropy gain are eliminated. Therefore, the features associated with internal nodes after pruning are considered as the most relevant features. Using the Matlab *treefit* routine, the decision tree based on samples extracted from [Tuncbag et al., 2009] showed that the 3 highest spectrum peak changes using EIIP, the energy change in the 7-th sub-band using EIIP and the global energy band using IC are the most appropriate candidates. These selected descriptors form a 5-dimensional vector called the sequence-based frequency-derived features in the sequel.

Structure-based features

The state-of-the-art 3D-structure-based features, namely, the solvent accessibility (accessible surface area (ASA)) [Lee and Richards, 1971], the pair potentials [Tuncbag et al., 2009] and the computational binding free energy change (Robetta) [Kortemme and Baker, 2002, Kortemme et al., 2004] are taken into account. In the recent work of Tuncbag *et al.* [Tuncbag et al., 2009], these features have been shown to be relevant for hotspot identification. The conservation score is not considered because it was not included in the best decision rule reported in [Tuncbag et al., 2009]. It should also be noted that the conservation score is seemingly not discriminating enough between hotspot and non-hotspot residues [Tuncbag et al., 2009, Cho et al., 2009].

Solvent accessibility The relative ASA in the complex state and the relative difference ASA between the complex and the monomer states of residue j are defined as in [Tuncbag et al., 2009]:

$$relCompASA^j = \frac{ASA^j_{comp}}{ASA^j_{max}} \times 100$$

$$relDiffASA^j = \frac{ASA^j_{mono} - ASA^j_{comp}}{ASA^j_{max}} \times 100$$

where ASA^j_{mono} (resp. ASA^j_{comp}) is the ASA of the j-th residue in monomer (resp. complex) state and ASA^j_{max} is its maximum ASA in a tri-peptide state.

Pair potentials The contact potential of residue j is defined as:

$$Potential^j = \text{abs}(\sum_{k=1}^{L} Pair(j,k))$$

where L is the number of residues and $Pair(j,k)$ is the contact potential of residues j and k. Two residues are considered to be in contact if they are closer than 7.0Å to each other in space and are separated by at least 3 residues in sequence [Tuncbag et al., 2009]. We thus have

$$Pair(j,k) = \begin{cases} p(j,k) & \text{if } d(j,k) \leqslant 7.0 \text{ and } |k-j| \geqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

in which $p(j,k)$ is the knowledge-based solvent-mediated potential [Keskin et al., 1998] between two residues at positions j and k, while $d(j,k)$ is the distance between their centers.

Computational binding free energy change (Robetta) These values, given by the Robetta server [Kortemme et al., 2004], are changes in computational binding free energy. The calculation is based on the energy function, proposed in [Kortemme and Baker, 2002], which takes into account Lennard-Jones potential, hydrogen bonding and solvation interaction.

The first three structure-based features can be retrieved through the HOTPOINT server [Tuncbag et al., 2010] and the fourth one from the Robetta server [Kortemme et al., 2004].

2.3.2 Dataset

The evaluation is performed on the union of ground-truth datasets considered in recent works [Tuncbag et al., 2009, Cho et al., 2009] dedicated to hotspot detection.

In this union, we consider only the experimental alanine scanning data with available measured values of $\Delta\Delta G$. These data were extracted by Tuncbag [Tuncbag et al., 2009] and Cho [Cho et al., 2009] from the ASEdb [Thorn and Bogan, 2001] and the published dataset of [Kortemme and Baker, 2002], after removing redundancy that could bias the training and/or the classification performance measurements. More specifically, they excluded homologous proteins with more than 35% sequence identity. Furthermore, in [Cho et al., 2009], proteins with high structural similarity (structure alignment score is higher than 80) were also discarded. This ensures to consider proteins from different families and avoids potential bias due to the over-representation of a given protein type. Data from BID (Binding Interface Database) [Fischer et al., 2003] are not included because they do not provide the measured values of the change in binding free energy ($\Delta\Delta G$). The resulting dataset involves an important variability in terms of protein sequences and, as such, provides a relevant basis for the evaluation of hotspot detection techniques.

To label the residues of the dataset, we proceed as in [Tuncbag et al., 2009]. Specifically, residues associated with a value of $\Delta\Delta G$ greater than or equal to 2.0 kcal/mol when mutated to alanines are deemed as hotspots and those with $\Delta\Delta G$ less than 0.4 kcal/mol are regarded as non-hotspots. The other residues are not included in the dataset in order to better discriminate the two classes. The final two-class dataset[1] contains 221 residues in which 76 are hotspots and 145 are non-hotspots. This dataset is somewhat unbalanced with the hotspot class representing only 34% of the samples. The amino acid chains considered in the dataset are listed in Table 2.2. The detailed information on these sequences can be obtained from the Protein Data Bank (PDB) [Berman et al., 2000] via their entry identities (PDB ids) and chain identities (Chain ids).

2.3.3 Hotspot identification performance assessment results

To assess the identification performance, we consider six usual evaluation measures: Accuracy (A), Precision (P), Recall (R), Specificity (Sp), F-measure ($F1$) and Matthews correlation coefficient (MCC). These measures are defined as follows:

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$Sp = \frac{TN}{TN + FP}$$

[1] The dataset is available at http://perso.telecom-bretagne.eu/pastor/.

Table 2.2 — Amino acid chains present in dataset

PDB id	Chain id	Molecule
1a4y	A	RNase inhibitor
	B	Angiogenin
1ahw	C	Tissue factor
1brs	A	Barnase
	D	Barstar
1bxi	A	Colicin E9 immunity protein
1cbw	D	BPTI Trypsin inhibitor
1dan	L	Blood coagulation factor VIIA
	T, U	Soluble tissue factor
1dvf	A, B	FV D1.3
1f47	A	Cell division protein FTSZ
1fc2	C	Fragment B of protein A complex
1fcc	C	Streptococcal protein G (C2 fragment)
1gc1	C	CD4
1jrh	I	Interferon-gamma receptor alpha chain
1jtg	A	Beta-lactamase tem
	B	Beta-lactamase inhibitory protein
1nmb	L	FAB NC10
1vfb	A	IGG1-KAPPA D1.3 FV (light chain)
	B	IGG1-KAPPA D1.3 FV (heavy chain)
	C	Hen egg white lysozyme
2ptc	I	Trypsin inhibitor
3hfm	H	HYHEL-10 IGG1 FAB (heavy chain)
	L	HYHEL-10 IGG1 FAB (light chain)
	Y	Hen egg white lysozyme
3hhr	A	Human growth hormone
	B	Human growth hormone receptor (hGHbp)

$$F1 = 2 \times \frac{P \times R}{P + R}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}}$$

where: TP (resp. TN) is the number of true positives (resp. true negatives), defined as the number of samples that are correctly predicted as hotspots (resp. non-hotspots); FP (false positive) is the number of non-hotspots that are falsely predicted as hotspots, and FN (false negative) is the number of hotspots that are not detected.

Because of the unavoidable trade-off between precision and recall on the one hand, and between recall and specificity on the other hand, both $F1$ and MCC are very usual in machine learning as quality measures of binary classification. The F-measure ($F1$) balances precision P and recall R only, whereas the Matthews correlation coefficient (MCC) takes into account the four terms TP, TN, FP, FN of the confusion matrix. Let us note that a predictor should not perform worse than the 'random guess', 'all-

are-positives' and 'all-are-negatives' ones. Therefore, it should satisfy the following conditions:

$$MCC > MCC_{rand}$$

$$F1 > \max(F1_{rand}, F1_{pos}, F1_{neg})$$

where: MCC_{rand} and $F1_{rand}$ are expected values of MCC and $F1$ scores for the 'random guess' predictor; $F1_{pos}$ and $F1_{neg}$ are F-measure values for the 'all-are-positives' and the 'all-are-negatives' predictors, respectively. In case of a dataset with p positives and n negatives, these conditions can easily be proved to become $MCC > 0$ and $F1 > 2p/(2p+n)$. With a simple calculation, the significant thresholds of MCC and $F1$ can be found to be $MCC_{thres} = 0$ and $F1_{thres} = 0.51$ for our evaluation dataset of 76 hotspots and 145 non-hotspots.

To demonstrate the relevance of our sequence-based frequency-derived features (1DFreq), we compare their predictive performance with that of 3D structure-based ones (3DStruct), i.e. relCompASA, relDiffASA, Potential and Robetta, in terms of the six aforementioned measures, especially $F1$ and MCC. The results are also compared with those obtained using the empirical rule introduced by Tuncbag in HOTPOINT [Tuncbag et al., 2010], which is shown in [Tuncbag et al., 2009] to provide similar results to Robetta [Kortemme et al., 2004] and outperform other state-of-the-art methods including KFC (Knowledge-based FADE and Contacts) [Darnell et al., 2007]. This empirical model only requires two out of the four 3D structure-based features to achieve hotspot recognition:

$$isHotspot = (relCompASA \leqslant 20\%) \text{AND} (Potential \geqslant 18.0)$$

In the sequel, for the sake of convenience, the group of these two features, i.e. relCompASA and Potential, will be referred to as 3DHotpoint. The recognition results obtained by combining structure-based features (3DStruct or 3DHotpoint) with our sequence-based 1DFreq are also presented.

The quantitative evaluation is carried out through repeated 10-fold cross-validations. One single splitting of the original dataset into a training set and a test set may yield an over/under estimation of the recognition performance. In contrast, the cross-validation procedure involves multiple splittings and avoids such over/under estimation. Specifically, in a 10-fold cross-validation, the dataset is first randomly partitioned into 10 mutually exclusive subsets (or folds) of nearly equal size (about 22 residues). This partition is processed in such a way that all folds contain approximately the same proportion of hotspots and non-hotspots as the original dataset. By such a stratified sampling, each fold is a good representative of the whole dataset. Given a partition, 10 training-testing iterations are subsequently performed. In each

Table 2.3 — Classification performance results (mean(±standard deviation))

Features	3DHotpoint[a,d]	3DStruct[b]	1DFreq[c]	3DHotpoint+1DFreq	3DStruct+1DFreq
Accuracy (A)	0.729	0.751(±0.010)	0.790(±0.013)	0.798(±0.014)	0.824(±0.009)
Precision (P)	0.629	0.672(±0.021)	0.748(±0.025)	0.751(±0.025)	0.801(±0.017)
Recall (R)	0.513	0.541(±0.018)	0.589(±0.029)	0.616(±0.031)	0.649(±0.017)
Specificity (Sp)	0.841	0.861(±0.012)	0.896(±0.014)	0.893(±0.013)	0.915(±0.009)
$F1$	0.565	0.599(±0.016)	0.659(±0.023)	0.676(±0.025)	0.716(±0.015)
MCC	0.375	0.427(±0.024)	0.518(±0.031)	0.537(±0.033)	0.597(±0.020)

[a] 3DHotpoint: relCompASA and Potential.
[b] 3DStruct: relCompASA, relDiffASA, Potential and Robetta.
[c] 1DFreq: our proposed sequence-based frequency-derived features.
[d] The results presented in this column are obtained by HOTPOINT while others are yielded by using RF with $nbTrees$= 1000 classification trees. For RF, all possible values of $mTry$ are tested and the best results are provided.

Table 2.4 — Results given by different t-tests [a]

Null hypothesis (H_0)	Alternative hypothesis (H_1)	$F1$		MCC	
		Accept	p-value	Accept	p-value
1DFreq $\leqslant$ 3DHotpoint [b]	1DFreq $>$ 3DHotpoint	H_1	2.62×10^{-63}	H_1	2.08×10^{-69}
1DFreq $\leqslant$ 3DStruct	1DFreq $>$ 3DStruct	H_1	5.04×10^{-50}	H_1	7.70×10^{-58}
[3DHotpoint+1DFreq] $\leqslant$ 1DFreq	[3DHotpoint+1DFreq] $>$ 1DFreq	H_1	2.89×10^{-07}	H_1	2.60×10^{-05}
[3DStruct+1DFreq] $\leqslant$ 1DFreq	[3DStruct+1DFreq] $>$ 1DFreq	H_1	1.21×10^{-48}	H_1	1.37×10^{-50}

[a] Right-side t-tests were performed. In this table, the notation FeasA $>$ FeasB (resp. FeasA $\leqslant$ FeasB) denotes the hypothesis that the mean performance score provided by FeasA is greater than (resp. less than or equal to) that yielded by FeasB.
[b] The results reported in this row are obtained using one-sample t-test while others are provided by two-sample ones.

iteration, a different fold (22 residues) is taken as the test set and the remaining 9 folds (199 residues) serve as the training set. The results from the 10 iterations are then grouped to evaluate the classification performance. The 10-fold cross-validation is repeated multiple times (with different stratified partitions) to assess the variability of the performance measures. In this study, a 100×10-fold cross-validation is used. The results[2] for the considered dataset are reported in Table 2.3. The prediction performance of HOTPOINT for the same dataset is also presented for reference. In Figure 2.3, the boxplots of $F1$ and MCC scores yielded by different groups of features are included for better comparison. The statistical significance of the results is further assessed by examining the p-values obtained using Student's t-tests. The statistical significance level is set to $\alpha = 0.01$. Table 2.4 provides the results of different t-tests obtained using the Matlab Statistics Toolbox.

[2]The Matlab code used to obtain these results is available at http://perso.telecom-bretagne.eu/pastor/.

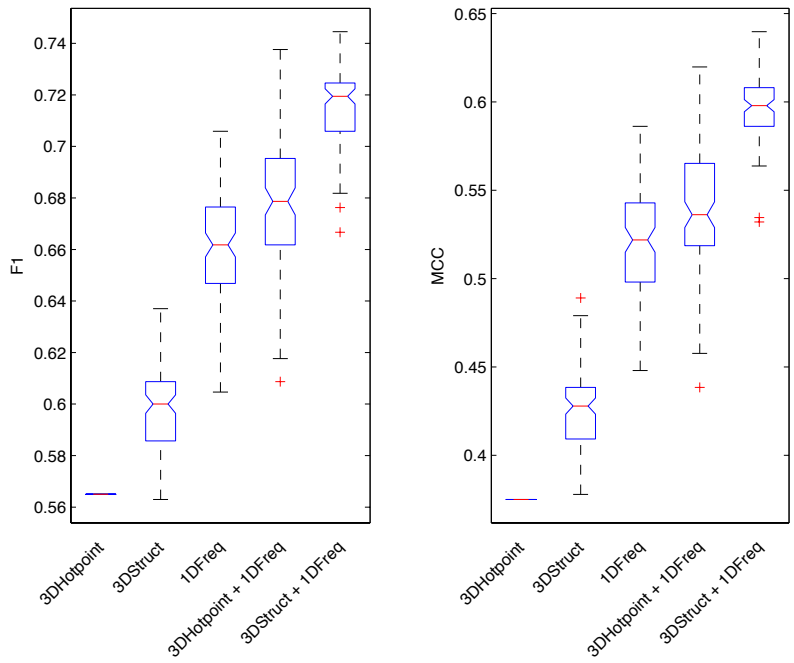


Figure 2.3 — Boxplots of $F1$ (left) and MCC (right) score values yielded by different sets of features. These boxplots were obtained using the Matlab *boxplot* routine with the default parameters. For a given boxplot, the extremes of the triangular notch represent the endpoints of the so-called comparison interval of the median at the 5% significance level. Two medians are considered to be significantly different if their comparison intervals do not overlap. The red crosses denotes the outliers.

2.4 Discussion

2.4.1 Relevance of sequence-based frequency-derived features with respect to previous work

Sequence-based descriptors can predict hotspots

The reported quantitative evaluation demonstrates the relevance of the proposed frequency-based protein sequence features for hotspot recognition compared to previous work. The results show that 1DFreq reach better detection scores than 3DStruct with respect to all six performance measures. More hotspots are detected (59% compared to 54%) and they are detected more precisely (75% compared to 67%). These features then yield higher $F1$ and MCC scores than 3DStruct (0.66 and 0.52 compared to 0.60 and 0.43 respectively). The t-tests on $F1$ and MCC stress the statistical significance of this improvement with p-values ≈ 0 ($\ll 0.01$). With an $F1$ score of 0.66 (>0.51) and an MCC score of 0.52 (>0), the hotspot recognition based on the proposed protein sequence features is meaningful.

Besides the relevance for hotspot recognition, an additional advantage of the pro-

posed approach with respect to previous works, especially those relying on features extracted from the 3D protein structure, is its low complexity. It only relies on the analysis of the numerical representations of the 1D sequence of amino acid using frequency analysis. By contrast, the reconstruction of the 3D structure of a protein is a complex task requiring complex experimental expertise, especially regarding protein crystallization to achieve a 3D imaging of the protein structure. Such crystallization issues are particularly complex for large compounds [Alberts et al., 2010]. Relying solely on the 1D sequence, we enlarge the potential application field of hotspot recognition techniques, especially for newly-sequenced proteins presenting weak homologies to proteins with known 3D structures [Rost, 1999, Pandini et al., 2007].

From an engineering point of view, the proposed approach is much less complex than those based on docking, MD simulations, graph analysis or 3D structure information derived descriptors. Moreover, by construction, the Random Forest algorithm is highly parallelizable and can easily be implemented on low cost hardwares such as graphic processors (GPUs). As a result, our method should be capable of dealing with large-scale datasets, which become a crucial problem as more and more proteomic data are available in the public domain [Vaidyanathan and Yoon, 2004, Deergha Rao and Swamy, 2008].

The combination of 3D structure characteristics and 1D frequency-based features improves the recognition of hotspots

We also evaluated the combination of the proposed 1D sequence features and descriptors of the 3D structure. As reported in Table 2.3, the combination [3DStruct+1DFreq] leads to significant recognition statistics (p-values < 0.01) with an accuracy of 82% and a precision of 80%. It is proved to reach better recognition performance than the 1D sequence features (i.e. 1DFreq) alone or the combination [3DHotpoint+1DFreq] (respectively, 82% vs. 79% and 80% for recognition accuracy and 80% vs. 75% and 75% for recognition precision). It is also worth noticing that [3DStruct+1DFreq] returns a significant gain for all six assessment indices.

These results show that the proposed frequency-based 1D sequence features provide discriminative information complementary to the descriptors issued from the classical local characteristics of the 3D structure of the protein. It then provides the means to improve recognition performance for a subset of protein sequences whose 3D structures are known. It may also provide the basis for similar improvements for protein sequences having high homology (typically, greater than 35% of residue identity) with a protein whose 3D structure is known. For such a homology level, it is indeed generally assumed that the 3D structure of the analyzed protein can be inferred from its homologue [Rost, 1999]. One may expect that the combination of the proposed 1D sequence features and

of 3D features extracted from the inferred structure could also lead to substantial improvement of hotspot recognition compared to 1D sequence features alone.

2.4.2 Physico-chemical interpretation of the proposed features

The analysis of frequency-based features of 1D numerical representations of the protein amino acid sequence was initially motivated by the RRM [Cosic, 1994], a physico-mathematical model which was originally introduced as an attempt to get an insight into the selectivity of protein interactions. By assigning to each amino acid a physical parameter value relevant to the protein bioactivity and analyzing the resulting numerical sequence, the RRM has successfully revealed the existence of frequency characteristics that characterize how a protein can recognize its target in an interaction. From the RRM perspective, proteins of the same family, sharing the same biological function, also share some frequency-based features. In particular, their frequency spectra exhibit a common *characteristic frequency* [Cosic, 1997]. This characteristic frequency was identified from the consensus spectrum, which is defined as the multiple cross-spectrum function of the Fourier transforms of all the sequences of the protein family as in [Cosic, 1997]:

$$M(n) = |X_1(n)|.|X_2(n)|...|X_K(n)|, n = 0, 1, ..., N-1$$

where $X_i(n), i = 1, 2, ..., K$ are the discrete Fourier transform coefficients of the numerical representation of the i-th protein sequence of the family, K is the number of family sequences and N is the length of the longest sequence. Shorter sequences are filled up with their mean value to have the same length N. Figure 2.4 reports the consensus spectrum of the fibroblast growth factor (FGF) family. This consensus spectrum clearly exhibits a characteristic frequency at $f_c = 0.4567$, which is significantly present in all the sequences of the FGF family.

It was conjectured in [Cosic, 1997] that these characteristic frequencies are associated with the common function of the proteins of a given family. Since hotspots are referred to as the key positions that determine the protein function, they were defined by Cosic et al. [Cosic, 1997] as the residues that are most affected by any change made to the amplitude spectrum at the characteristic frequency corresponding to the protein biological function. Although some evidence of the correlation between the hotspots defined by RRM and those detected by ASM were reported [Ramachandran et al., 2004, Ramachandran and Antoniou, 2008, Sahu and Panda, 2009], the recognition performance was limited to very few examples. Besides, earlier applications of the RRM required the functional family of the protein to be known to compute the corresponding characteristic frequency. Our approach does not impose such a constraint. Rather than

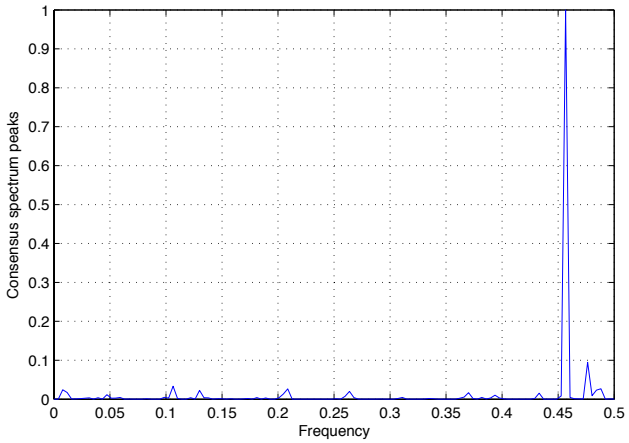


Figure 2.4 — The characteristic frequency of the fibroblast growth factor (FGF) protein family: the consensus Fourier spectrum shows that the FGF protein family members share a common characteristic frequency at $f_c = 0.4567$. In this example, the EIIP values have been used.

a purely DSP-based approach aimed at detecting local residues associated with the characteristic frequency, we combine DSP tools and mutagenesis principles. We locally determine frequency-related energy changes resulting from the computational mutation of residue subsets to alanines. Considering the alanine mutations as a reference model, our procedure can be applied to newly sequenced or unclassified proteins, which might enlarge its potential application domain. Moreover, we have reported an actual evaluation of hotspot recognition performance with respect to a reference database of experimental ASM hotspots, which is, to the best of our knowledge, the largest available validated dataset of hotspots.

Our results bring new evidence to support the conjecture of Cosic et al. [Cosic, 1997] that protein hotspots are associated with frequency features of physico-chemical characteristics of the amino acid sequence. Whereas this statement was analyzed in [Cosic, 1997] for the RRM model associated with electron-ion interaction potentials, we have shown here that protein hotspots may also involve specific frequency-related features for other physico-chemical characteristics such as ionization constants. Future work should further investigate, from both the computational and the biophysical point of view, the characterization and the interpretation of such frequency-related properties of protein and associated hotspots.

2.4.3 Comparison to other DSP-based hotspot detection methods

As aforementioned, motivated by the finding of the protein characteristic frequency in [Cosic, 1994], many studies, such as those in [Ramachandran et al., 2004, Ramachan-

dran and Antoniou, 2008, Deergha Rao and Swamy, 2008, Sahu and Panda, 2009], have been carried out to solve the hotspot detection problem by digital signal processing (DSP) techniques. Basically, by analyzing the signal representing the considered amino acid sequence in the transform domain, these approaches attempt to locate the portion of the equivalent signal that contributes the most to the characteristic frequency, and by thus, to identify hotspots. For example, in [Ramachandran et al., 2004], a Short-time Fourier transform was used and high-energy regions in the time-frequency spectrum were investigated. Similarly, in [Deergha Rao and Swamy, 2008], the wavelet transform was considered and in [Sahu and Panda, 2009], the S-Transform was taken into account. Although being illustrated via a few well-known protein families, the detection results presented by these approaches were somehow limited. With regard to the dimension of the problem, such methods based on a single descriptor, which characterizes specific high-energy regions in the transform domain, can hardly provide a good solution in practice. In this respect, the approach proposed in this work has overcome this limitation by making it possible to get multiple descriptors involved. These descriptors can be of various nature and can be resulted from different measurements and processings in practice. The descriptors yielded by the transformations referred in this section could also be included. In this respect, the machine learning based method exposed in this chapter presents relevant detection results.

2.4.4 Future work

In this application, an *in-silico* alanine scanning framework with frequency-derived features of numerical representations of the amino acid sequences has been introduced for protein hotspot recognition. It outperforms previous work on a ground-truth database of protein hotspots [Tuncbag et al., 2009, Cho et al., 2009]. We have also shown that improved recognition performance can be achieved when the 3D structure of the protein is available, i.e. from the combination of the proposed 1D frequency-related features and local descriptors of the 3D structure. The reported experiments support the assumption that the functionality of a protein is basically encoded into its primary amino acid sequence. But how this encoding is performed is still an open question. In this respect, it could be profitable to get a better insight into the physico-chemical meaning of the frequency-related descriptors introduced in this work.

As mentioned in Section 2.2.2, other time series analyses can be involved in the proposed framework to provide new hotspot descriptors. The use of DSP techniques such as those in [Ramachandran et al., 2004, Ramachandran and Antoniou, 2008, Deergha Rao and Swamy, 2008, Sahu and Panda, 2009] might be investigated to derive descriptors that could further be compared to and/or combined with ours for proteins belonging to the same functional family.

The main focus of this chapter was not the classifier itself, but rather the relevance — assessed by classification performance measurements — of the proposed descriptors as hotspots signatures. Therefore, it can be expected that the classification performance could perhaps be even further improved by combining RF with other classifiers such as SVM, neural networks and so forth. An exhaustive study of this type could be addressed in another work.

Conclusion

As non-parametric approaches, machine learning brings robustness against any variation of observations. Information on the conditional distribution of observations under each permissible hypothesis — in classification, it is the observation distribution of each class — is completely unknown prior to either the training phase or the predicting phase of the predictor. As such, the predictor input— i.e. descriptors — may vary in as many ways as one can imagine. However, the most discriminants descriptors, which best separate the distributions of the classes, should be chosen to obtain good prediction performance. This somehow requires some prior knowledge on how the difference in predicted classes may reflect the difference in the considered observations. Moreover, in some cases, this prior knowledge might be required to better parameterize the classifier. For example, the choice of distance/divergence in kernel based methods usually depends on how the classes are self-organized in the feature space.

To compensate the lack of knowledge on the conditional distributions of observations under the permissible hypotheses, the machine learning approach performs a training phase, which, in turn, requires a training set of data instance. This training dataset must be sufficiently large and representative, allowing the principle characteristics of the observation distributions to be learned by the predictor through the training process. The construction of such a dataset is sometimes laborious in many applications. The detection of protein interaction hotspots presented in this work is a typical example. The finding of 3D structure based descriptors and the annotation of the nature (hotspot/non-hotspot) of a residue require many experiments, such as protein crystallization, imaging, measuring of binding free-energy, etc. Therefore, the constituted dataset is somehow limited. In particular, a hotspot dataset with 221 individuals might be sufficient for a preliminary study as in this work. It is yet too modest for a better and more complete study to be done. More data are then required.

Additionally, for such a prediction in the Machine Learning framework, the best assessment that can be provided is the estimation of the prediction performance carried out on a test dataset. Ideally, this test set is an independent dataset that exhibits the same observation distributions as those in the training set to be used. No optimality on the performance of the trained predictor can be assured nor any meaningful per-

formance boundaries can be suggested for an arbitrary practical dataset. Therefore, in practice, the actual prediction performance might be very different from that given by the estimation. This still presents a limitation of learning-based approach in practice, especially in applications in which some boundaries on the prediction performance must be guaranteed, such as: medical application, structure health monitoring, etc.

In these respects, other non-parametric approaches, which take into account as little prior information on the observations as possible, yet provide some performance optimality, can be sought. The RDT (Random Distortion Testing) framework is among such approaches. In the second part of this thesis, this framework will be presented with illustrations via a proposed automatic monitoring framework in medical domain.

Part II

Detection in the Random Distortion Testing framework and application to mechanical ventilation system monitoring

CHAPTER 3

Random Distortion Testing and Signal detection

With respect to the aforementioned limitations of learning-based framework, we can propose a new statistical decision strategy, which brings both robustness and statistical optimality at the same time. The Random Distortion Testing (RDT), recently introduced by D. Pastor in [Pastor, 2011, Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a], provides such a decision approach. This approach is motivated by many practical applications in which it is required to evaluate — on the basis of random observation in noise — the corruption of the signal of interest with respect to some known model. These applications include radar, sonar, remote sensing, telecommunication, etc.

To begin with, let us consider the observation vector $\mathbf{Y}$ captured by a sensor. $\mathbf{Y}$ is supposed to be in additive noise $\mathbf{X}$, i.e.:

$$\mathbf{Y} = \boldsymbol{\Theta} + \mathbf{X} \tag{3.1}$$

where the d-dimensional vector $\boldsymbol{\Theta}$ is the signal of interest. Very often, $\boldsymbol{\Theta}$ is actually random with unknown distribution. Given some nominal deterministic model $\boldsymbol{\theta}_0$, it is then of interest to verify whether or not a realization $\boldsymbol{\theta}$ of $\boldsymbol{\Theta}$ is a corrupted version of $\boldsymbol{\theta}_0$. This problem is usually expressed as the testing of the simple null hypothesis $[h_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0]$ against the composite alternative one $[h_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0]$. As announced in [Lehmann and Romano, 2005, Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a], no Uniformly Most Powerful (UMP) test exists for this problem. In the deterministic case where the signal of interest $\boldsymbol{\Theta}$ is not a random vector but an unknown deterministic vector $\boldsymbol{\theta}$, the so-called *holy trinity* — i.e. the generalized likelihood ratio test (GLRT) [Neyman and Pearson, 1928], the Rao score test [Rao, 1948] and the Wald test [Wald, 1943] — could provide powerful test as long as a sufficient number of independent observations can be collected to benefit from the asymptotic properties of the maximum likelihood estimates and Fisher's information matrix. The general case with random signal $\boldsymbol{\Theta}$ is still an open problem, which the RDT proposed in [Pastor, 2011, Pastor

and Nguyen, 2012b, Pastor and Nguyen, 2012a] attempts to investigate.

It should be noted that, in real-world applications, due to unavoidable unknown random fluctuations of environment regardless of noise, small signal distortion from the nominal model is practically of poor interest. Therefore, testing $[\boldsymbol{\theta} = \boldsymbol{\theta}_0]$ might be too strict — and even impossible — because of physics. It is then more reasonable to introduce some value τ to tolerate possibly small distortions around $\boldsymbol{\theta}_0$ via testing the composite hypothesis that the signal still remains in a neighborhood of $\boldsymbol{\theta}_0$, specified by tolerance τ. Let $\mathbf{X}$ be independent Gaussian noise with known positive-definite covariance matrix C. In other words, the signal is supposed to be observed in additive, independent Gaussian noise, a reasonable model in many practical applications. By compensating the variation induced by the noise covariance matrix C, the Mahalanobis norm [Mahalanobis, 1936] of $\boldsymbol{\Theta} - \boldsymbol{\theta}_0$ is suitable for assessing how much the signal $\boldsymbol{\Theta}$ deviates from its nominal model $\boldsymbol{\theta}_0$. The test is then to verify whether or not the Mahalanobis norm of $\boldsymbol{\Theta} - \boldsymbol{\theta}_0$ exceeds the specified non-negative real-valued tolerance τ, i.e. testing the null event or hypothesis $[h_0 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant \tau]$ against its alternative one $[h_1 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau]$, where $\|.\|$ is the Mahalanobis norm. Such a testing problem is named Random Distortion Testing (RDT). No information on the observation distribution is required. No training database is needed. The test relies exclusively on knowledge of the observation noise, which is possibly estimated in practice. This chapter begins with some preliminary material, providing basic information for understanding RDT. Main results in both the deterministic and random cases are then briefly presented. Finally, the application of RDT in signal detection is carried out with an example illustrating the properties of RDT with respect to those given by classical testing frameworks.

3.1 Preliminary material

In what follows, all random vectors and variables are defined on the same probability space $(\Omega, \mathcal{B}, \mathrm{P})$. Using the same notation as in [Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a], $\mathcal{M}(\Omega, \mathbb{R}^d)$ denotes the set of all d-dimensional random real vectors defined on $(\Omega, \mathcal{B})$. Given any random vector $\mathbf{Z} \in \mathcal{M}(\Omega, \mathbb{R}^d)$, $\mathrm{P}_{\mathbf{Z}}$ represents the probability distribution of $\mathbf{Z}$, i.e. the probability measure defined by $\mathrm{P}_{\mathbf{Z}}(B) = \mathrm{P}[\mathbf{Z} \in B]$ for any Borel subset B of $\mathbb{R}^d$.

The Mahalanobis norm

Given $d \times d$-dimensional positive-definite covariance matrix C of noise, the Mahalanobis norm $\|.\|$ in $\mathbb{R}^d$ is defined by:

$$\|\mathbf{y}\| = \sqrt{\mathbf{y}^T \mathrm{C}^{-1} \mathbf{y}}, \forall \mathbf{y} \in \mathbb{R}^d, \tag{3.2}$$

where A^T is the transpose of any matrix or vector A. It should be noted that the positive-definite covariance matrix C can be eigen-decomposed as $\mathrm{C} = \mathrm{U}\Delta\mathrm{U}^T$, where diagonal matrix Δ contains all the eigenvalues of C and matrix U is constructed by the corresponding eigenvectors. By such a decomposition, the Mahalanobis norm of any random vector $\mathbf{y} \in \mathbb{R}^d$ can be replaced by the Euclidean norm of a corresponding vector in $\mathbb{R}^d$

$$\|\mathbf{y}\| = \|\Phi\mathbf{y}\|_2 \tag{3.3}$$

where $\Phi = \Delta^{-1/2}\mathrm{U}^T$ and $\|\mathbf{y}\|_2 = \sqrt{\mathbf{y}^T\mathbf{y}}$ is the standard Euclidean norm of $\mathbf{y}$.

The power function, the size and the power of a test

A *test* is a map from $\mathbb{R}^d$ into the set of indices $\{0, 1\}$. The returned value of such map indicates the index of the hypothesis that is accepted by the test. Given a test $\mathcal{T}$ on the basis of observation $\mathbf{Y}$ whose distribution belongs to the family $\mathcal{P} = \{\mathcal{N}(\boldsymbol{\theta}, \mathrm{C}) : \boldsymbol{\theta} \in \mathbb{R}^d\}$ of normal distributions with known variance covariance matrix C, as in [Lehmann and Romano, 2005], the *power function* of $\mathcal{T}$ is the map that assigns to every $\boldsymbol{\theta} \in \mathbb{R}^d$ the value

$$\beta_{\boldsymbol{\theta}}(\mathcal{T}) = \mathrm{P}[\mathcal{T}(\mathbf{Y}) = 1], \tag{3.4}$$

which is the probability that $\mathcal{T}$ rejects the null hypothesis h_0, regardless of which hypothesis actually holds.

The *size* of $\mathcal{T}$ is defined as the least upper bound of the probability of false-alarm:

$$\alpha(\mathcal{T}) = \sup_{\boldsymbol{\theta} \text{ under } h_0} \beta_{\boldsymbol{\theta}}(\mathcal{T}) \tag{3.5}$$

The *power* of $\mathcal{T}$ is the value of the power function $\beta_{\boldsymbol{\theta}}(\mathcal{T})$ under the alternative hypothesis h_1, i.e. the detection probability.

Map $\mathcal{R}$

For an arbitrary non-negative value ρ, $\mathcal{R}(\rho, .)$ denotes the cumulative distribution function of the square root of any random variable following the non-central Chi-squared (χ^2) distribution with d degrees of freedom and non-central parameter ρ^2. It should be noticed that, for any $\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\theta}, \mathrm{C})$, we have $\Phi\mathbf{Z} \sim \mathcal{N}(\Phi\boldsymbol{\theta}, \mathbf{I}_d)$, where the matrix Φ is defined as above and $\mathbf{I}_d$ is the $d \times d$-dimensional identity matrix, and therefore, the cumulative distribution function of $\|\mathbf{Z}\|$ is given by:

$$\mathrm{P}[\|\mathbf{Z}\| < \eta] = \mathcal{R}(\|\boldsymbol{\theta}\|, \eta) \tag{3.6}$$

Some properties of map $\mathcal{R}$ and results:

- For any $\rho \geqslant 0$, $\mathcal{R}(\rho, .)$ is continuous and strictly increasing. As a result, it is a one-to-one mapping from $[0, \infty)$ to $[0, 1)$.
- For any $\eta > 0$, $\mathcal{R}(., \eta)$ is a strictly decreasing map.
- Given a value γ that $0 < \gamma \leqslant 1$ and $\rho \geqslant 0$, the equation $1 - \mathcal{R}(\rho, \eta) = \gamma$ admits a unique solution $\lambda_\gamma(\rho) > 0$ in η.
- Given $0 < \gamma \leqslant 1$, the aforementioned solution $\lambda_\gamma(.)$ is an everywhere continuous and strictly increasing map from $[0, \infty)$ to $[0, \infty)$.
- Given $\rho > 0$, $\gamma \mapsto \lambda_\gamma(\rho)$ is a continuous and strictly decreasing map from $(0, 1]$ to $[0, \infty)$.

The proof for these properties and results can be found in [Pastor and Nguyen, 2012b].

Thresholding tests on the Mahalanobis distance to the nominal model

Given a nominal model $\boldsymbol{\theta}_0$ for signal $\boldsymbol{\Theta}$ and a specified value $\eta \leqslant 0$, a *thresholding test* with *threshold* η on the Mahalanobis distance to $\boldsymbol{\theta}_0$ is any test $\mathcal{T}_\eta$ that

$$\mathcal{T}_\eta(\mathbf{y}) = \begin{cases} 1 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| > \eta \\ 0 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| \leqslant \eta \end{cases} \tag{3.7}$$

for any $\mathbf{y} \in \mathbb{R}^d$. The handling of equality plays no important role thanks to the absolute continuity of the probability distribution function of the observation $\boldsymbol{\Theta}+\mathbf{X}$ with respect to Lebesgue's measure in $\mathbb{R}^d$. The notion of randomized test also becomes useless.

3.2 Distortion Testing

3.2.1 Deterministic case (DDT)

With respect to the observation model in (3.1), the signal $\boldsymbol{\Theta}$ is now supposed to be an unknown deterministic vector $\boldsymbol{\theta}$. The observation $\mathbf{Y}$ is then normally distributed with the positive-definite covariance matric C and unknown mean $\boldsymbol{\theta} \in \mathbb{R}^d$, i.e. $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\theta}, \mathrm{C})$. Given a specific tolerance τ, the deterministic distortion testing (DDT) problem is the testing of the hypothesis $[h_0 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leqslant \tau]$ against the alternative $[h_1 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| > \tau]$ based on the observation in noise as follows:

$$\begin{cases} \text{Observation:} & \mathbf{Y} \sim \mathcal{N}(\boldsymbol{\theta}, \mathrm{C}) \\ \text{Tested or null hypothesis } (h_0): & \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leqslant \tau \\ \text{Alternative hypothesis } (h_1): & \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| > \tau \end{cases} \tag{3.8}$$

The problem is invariant, in the sense given in [Lehmann and Romano, 2005, Chapter 12, Section 6.1], with respect to ellipsoids Υ_ρ defined by $\Upsilon_\rho = \{\mathbf{y} \in \mathbb{R}^d : \|\mathbf{y} - \boldsymbol{\theta}_0\| = \rho\}$. As such, an ellipsoid-UMPI (*ellipsoid-Uniformly Most Powerful Invariant*) test with an arbitrary level γ $(0 < \gamma < 1)$ — a UMP (*Uniformly Most Powerful*) test with level γ among those that are invariant with respect to any ellipsoid Υ_ρ — exists. This test is thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ with threshold height $\lambda_\gamma(\tau)$ on the Mahalanobis distance to the nominal model $\boldsymbol{\theta}_0$:

$$\mathcal{T}_{\lambda_\gamma(\tau)}(\mathbf{y}) = \begin{cases} 1 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| > \lambda_\gamma(\tau) \\ 0 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| \leqslant \lambda_\gamma(\tau) \end{cases} \tag{3.9}$$

where $\mathbf{y}$ is an instance of random observation vector $\mathbf{Y}$.

Actually, these thresholding tests $\mathcal{T}_{\lambda_\gamma(\tau)}$ offer a more general statistical optimality, the *maximal constant power* property, on ellipsoids Υ_ρ with $0 \leqslant \rho < \tau$. More specifically, given a non-negative real number τ and a value γ $(0 < \gamma < 1)$, any thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ has size γ and maximal constant power over the family of all ellipsoids Υ_ρ for the DDT problem in (3.8) with:

$$\beta_{\boldsymbol{\theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right) = 1 - \mathcal{R}(\rho, \lambda_\gamma(\tau)) \tag{3.10}$$

for any $\rho \geqslant 0$ and $\boldsymbol{\theta} \in \Upsilon_\rho$ (see also [Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a]). This optimal criterion means that, for any test $\mathcal{T}$ with level γ and constant power on a given ellipsoid Υ_ρ — i.e. $\forall \boldsymbol{\theta}, \boldsymbol{\theta}' \in \Upsilon_\rho, \beta_{\boldsymbol{\theta}}(\mathcal{T}) = \beta_{\boldsymbol{\theta}'}(\mathcal{T})$ — with $\rho > \tau$, the power of this test $\mathcal{T}$ on Υ_ρ would never exceed the constant power $\beta_{\boldsymbol{\theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right)$ given by $\mathcal{T}_{\lambda_\gamma(\tau)}$.

Moreover, the optimal test $\mathcal{T}_{\lambda_\gamma(\tau)}$ is also unbiased, i.e. $\beta_{\boldsymbol{\theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right) \geqslant \gamma$ for any value of γ that $0 \leqslant \gamma < 1$.

Remarks

1. It should be noted that the standard two-side problem of testing the mean of a normally distributed random vector — i.e. testing the hypothesis $[h_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0]$ against the alternative one $[h_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0]$ based on the observation $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\theta}, \mathrm{C})$ with positive-definite C — is a special case of the DDT problem stated in (3.8) in which $\tau = 0$. [Wald, 1943, Proposition III, p.450] is therefore a particular case of the results above.

2. *DDT from-below and DDT from-above:*
   In practice, there are cases where one would like to test the hypothesis $[h'_0 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| > \tau]$ against $[h'_1 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leqslant \tau]$. Such a problem — in which the two hypotheses are exchanged — is called DDT *from-below* to differentiate from the DDT *from-above* stated in (3.8). It turns out that the optimal test for such

from-below problem is the thresholding test $\mathcal{T}'_{\lambda_{1-\gamma}(\tau)}$ *from-below* $\lambda_{1-\gamma}(\tau)$ defined by:

$$\mathcal{T}'_{\lambda_{1-\gamma}(\tau)}(\mathbf{y}) = \begin{cases} 1 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| \leqslant \lambda_{1-\gamma}(\tau) \\ 0 & \text{if} \quad \|\mathbf{y} - \boldsymbol{\theta}_0\| > \lambda_{1-\gamma}(\tau) \end{cases} \tag{3.11}$$

This test also has size γ and maximal constant power over the family of all ellipsoids Υ_ρ for the DDT problem *from-below* τ, with constant power:

$$\beta_{\boldsymbol{\theta}}\left(\mathcal{T}'_{\lambda_{1-\gamma}(\tau)}\right) = \mathcal{R}(\rho, \lambda_{1-\gamma}(\tau)) \tag{3.12}$$

3. In the one-dimensional case, given observation $Y \sim \mathcal{N}(\theta, \sigma^2)$, there is no UMP test for testing $[h_0 : |\theta - \theta_0| \leqslant \tau]$ against $[h_1 : |\theta - \theta_0| > \tau]$ (the problem *from-above* τ). However, any test

$$\mathcal{T}_{\sigma\lambda_\gamma(\tau/\sigma)}(y) = \begin{cases} 1 & \text{if} \quad |y - \theta_0| > \sigma\lambda_\gamma(\tau/\sigma) \\ 0 & \text{if} \quad |y - \theta_0| \leqslant \sigma\lambda_\gamma(\tau/\sigma) \end{cases} \tag{3.13}$$

is UMPU (UMP-unbiased) with size γ for such problem. On the contrary, for the problem *from-below* τ of testing $[h'_0 : |\theta - \theta_0| > \tau]$ against $[h'_1 : |\theta - \theta_0| \leqslant \tau]$, any test

$$\mathcal{T}'_{\sigma\lambda_{1-\gamma}(\tau/\sigma)}(y) = \begin{cases} 1 & \text{if} \quad |y - \theta_0| \leqslant \sigma\lambda_{1-\gamma}(\tau/\sigma) \\ 0 & \text{if} \quad |y - \theta_0| > \sigma\lambda_{1-\gamma}(\tau/\sigma) \end{cases} \tag{3.14}$$

is UMP with specified size γ.

3.2.2 Random case (RDT)

The observation model in (3.1) is now considered in the general case where the signal $\boldsymbol{\Theta}$ of interest is random with unknown distribution and independent of normally distributed additive noise $\mathbf{X}$ with positive-definite covariance matrix C (i.e. $\mathbf{X} \sim \mathcal{N}(0, \mathrm{C})$). As in the deterministic case, given a nominal model $\boldsymbol{\theta}_0$ and a specific tolerance $\tau \geqslant 0$, one would like to verify whether the signal of interest is a strongly distorted version of the expected model — i.e. $\|\boldsymbol{\Theta}(\omega) - \boldsymbol{\theta}_0\| > \tau$ — or not, on the basis of the random observation $\mathbf{Y}(\omega)$, where $\omega \in \Omega$. In other words, we must test the null event $[h_0 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant \tau]$ against the alternative event $[h_1 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau]$. The term *event* is employed to make it clear that the test is carried out on realization of random vector or random variable, not on its parameters.

The random distortion testing (RDT) problem is then stated as follows:

$$\begin{cases} \text{Observation:} & \mathbf{Y} = \boldsymbol{\Theta} + \mathbf{X} \begin{cases} \mathbf{X} \sim \mathcal{N}(0, \mathrm{C}) \\ \boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^d) \\ \boldsymbol{\Theta} \text{ independent of } \mathbf{X} \end{cases} \\ \text{Tested or null event } (h_0): & \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant \tau \\ \text{Alternative event } (h_1): & \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau \end{cases} \tag{3.15}$$

This problem is then called the *event testing* problem *from-above* tolerance τ. In practice, by simply switching the two events of interest, the event testing problem *from-below* tolerance τ can also formulated with the null event $[h'_0 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau]$ and the alternative one $[h'_1 : \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant \tau]$. Instead of considering an observation whose distribution belongs to a parameterized family as in standard approaches, the RDT problem concerns any class of random observation $\mathbf{Y} = \boldsymbol{\Theta} + \mathbf{X}$, with unknown random signal $\boldsymbol{\Theta}$ having any unknown distribution in independent gaussian noise $\mathbf{X}$ (cf. [Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a]).

The RDT problem in (3.15) admits no UMP test with level $\gamma \in (0, 1)$. However, it could profit from the fact that the problem is invariant with respect to ellipsoids — i.e. as long as the realization of the signal $\boldsymbol{\Theta}$ lies on an ellipsoid Υ_ρ, the problem is unchanged — as in DDT. It is then expected to choose an optimal test among those verifying such invariance. The thresholding test on the Mahalanobis distance to the nominal model $\boldsymbol{\theta}_0$, once again, provides such optimality. It has been shown in [Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a] that any thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ — defined by:

$$\mathcal{T}_{\lambda_\gamma(\tau)}\left(\mathbf{Y}(\omega)\right) = \begin{cases} 1 & \text{if} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| > \lambda_\gamma(\tau) \\ 0 & \text{if} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| \leqslant \lambda_\gamma(\tau) \end{cases} \tag{3.16}$$

for any $\omega \in \Omega$ — has size γ and *maximal constant conditional power* over the family of all ellipsoids Υ_ρ for the RDT problem in (3.15). This *maximal constant conditional power* property means that: Given any $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^d)$ and for $\mathrm{P}_{\|\boldsymbol{\Theta}-\boldsymbol{\theta}_0\|}$ – almost every $\rho > \tau$, $\mathcal{T}_{\lambda_\gamma(\tau)}$ has *constant conditional power function* given $\boldsymbol{\Theta} \in \Upsilon_\rho$ — i.e. $\mathrm{P}\left[\mathcal{T}_{\lambda_\gamma(\tau)}(\mathbf{Y}) = 1 \,\middle|\, \boldsymbol{\Theta} \in \Upsilon_\rho\right] = \beta_{\boldsymbol{\theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right)$ for any $\boldsymbol{\theta} \in \Upsilon_\rho$ — and

$$\mathrm{P}\left[\mathcal{T}_{\lambda_\gamma(\tau)}(\mathbf{Y}) = 1 \,\middle|\, \boldsymbol{\Theta} \in \Upsilon_\rho\right] \geqslant \mathrm{P}\left[\mathcal{T}(\mathbf{Y}) = 1 \,\middle|\, \boldsymbol{\Theta} \in \Upsilon_\rho\right] \tag{3.17}$$

for any test $\mathcal{T}$ with level γ and constant conditional power function given $\boldsymbol{\Theta} \in \Upsilon_\rho$. The constant conditional power of $\mathcal{T}_{\lambda_\gamma(\tau)}$ is given by:

$$\mathrm{P}\left[\mathcal{T}_{\lambda_\gamma(\tau)}(\mathbf{Y}) = 1 \,\middle|\, \boldsymbol{\Theta} \in \Upsilon_\rho\right] = 1 - \mathcal{R}(\rho, \lambda_\gamma(\tau)) \tag{3.18}$$

for any given $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^d)$ and $\mathrm{P}_{\|\boldsymbol{\Theta}-\boldsymbol{\theta}_0\|}$ – almost every $\rho \geqslant 0$ (see also [Pastor and Nguyen, 2012b, Pastor and Nguyen, 2012a]). Moreover, for any $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^d)$ such that $\mathrm{P}\left[\|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| > \tau\right] \neq 0$, we have:

$$\beta_{\boldsymbol{\Theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right) \geqslant 1 - \mathcal{R}(\tau', \lambda_\gamma(\tau)) \tag{3.19}$$

where τ' is defined as:

$$\tau' = \sup\left\{t \in \mathbb{R} : t \geqslant \tau, \mathrm{P}\left[\tau < \|\boldsymbol{\Theta} - \boldsymbol{\theta}_0\| \leqslant t\right] = 0\right\}. \tag{3.20}$$

The test $\mathcal{T}_{\lambda_\gamma(\tau)}$ is also unbiased and UMP within the class of tests with level γ and invariant with respect to ellipsoids, i.e. $\beta_{\boldsymbol{\Theta}}\left(\mathcal{T}_{\lambda_\gamma(\tau)}\right) \geqslant \beta_{\boldsymbol{\Theta}}\left(\mathcal{T}\right)$ for any $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^d)$ and any ellipsoid-invariant test $\mathcal{T}$ with level γ.

For the RDT problem *from-below* tolerance τ, the test $\mathcal{T}'_{\lambda_{1-\gamma}(\tau)}$ *from-below* threshold $\lambda_{1-\gamma}(\tau)$ — defined by:

$$\mathcal{T}'_{\lambda_{1-\gamma}(\tau)}\left(\mathbf{Y}(\omega)\right) = \begin{cases} 1 & \text{if} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| \leqslant \lambda_{1-\gamma}(\tau) \\ 0 & \text{if} \quad \|\mathbf{Y}(\omega) - \boldsymbol{\theta}_0\| > \lambda_{1-\gamma}(\tau) \end{cases} \tag{3.21}$$

— is suggested. This test satisfies similar optimality as $\mathcal{T}_{\lambda_\gamma(\tau)}$ proposed for the RDT problem *from-above* tolerance τ.

3.3 Signal detection in RDT framework

In this section, the detection of a non-null unknown signal via its observation in noise, a problem of interest in many practical applications, is considered to illustrate the use of the RDT framework. The signal is supposed to be d-dimensional and noise is assumed additive independent gaussian. It will be shown that, with model mismatch, conventional approaches such as Neyman-Pearson's might fail, whereas the proposed RDT remains functioning in any case.

Let $\boldsymbol{\Xi}$ be the d-dimensional random signal of interest with unknown distribution such that $\boldsymbol{\Xi}$ is independent of the additive gaussian noise $\mathbf{X} \sim \mathcal{N}(0, \mathrm{C})$ with C positive-definite. Assuming that $\boldsymbol{\Xi} \neq 0$ (almost surely). The detection of $\boldsymbol{\Xi}$ is considered as the binary hypothesis testing problem where the null hypothesis H_0 is that only noise is present and the alternative one H_1 is that the signal of interest is along with additive noise. Since $\boldsymbol{\Xi} \neq 0$ (almost surely), there then exists a value $\tau' \geqslant 0$ such that $\|\boldsymbol{\Xi}\| > \tau'$ (almost surely). Let $\mathbf{Y}$ be the observation. The problem can be written as:

$$\begin{cases} H_0: & \mathbf{Y} \sim \mathcal{N}(0, \mathrm{C}) \\ H_1: & \mathbf{Y} = \boldsymbol{\Xi} + \mathbf{X} \text{ with } \mathbf{X} \sim \mathcal{N}(0, \mathrm{C}), \mathrm{P}\left[\|\boldsymbol{\Xi}\| > \tau'\right] = 1 \end{cases} \tag{3.22}$$

To cast this detection problem in the RDT theoretical framework, we introduce the random variable ε whose value, taken from the set $\{0, 1\}$, represents the index of the true hypothesis — i.e. indicates the presence (when $\varepsilon = 1$) or the absence (when $\varepsilon = 0$) of the signal $\boldsymbol{\Xi}$ in observation $\mathbf{Y}$. This random variable ε is independent of signal $\boldsymbol{\Xi}$ and noise $\mathbf{X}$. The observation can be rewritten as: $\mathbf{Y} = \varepsilon\boldsymbol{\Xi} + \mathbf{X}$. The problem then amounts to testing $[\|\varepsilon\boldsymbol{\Xi}\| = 0]$ against $[\|\varepsilon\boldsymbol{\Xi}\| > \tau']$. This is obviously the RDT problem (3.15) with $\boldsymbol{\Theta} = \varepsilon\boldsymbol{\Xi}$, $\boldsymbol{\theta}_0 = 0$ and $\tau = 0$ and the optimal test proposed by RDT is then $\mathcal{T}_{\lambda_\gamma(0)}$. The size of such a test — which now, turns out to be the probability of false-alarm of the signal detector — is given by:

$$\mathrm{P}_{\mathrm{fa}}\left[\mathcal{T}_{\lambda_\gamma(0)}\right] = \mathrm{P}\left[\|\mathbf{X}\| > \lambda_\gamma(0)\right] = 1 - \mathcal{R}\left(0, \lambda_\gamma(0)\right) = \gamma \tag{3.23}$$

and the power of the test — which is now the detection probability — is lower-bounded as follows:

$$\mathrm{P_d}\left[\mathcal{T}_{\lambda_\gamma(0)}\right] = \int_{\tau'}^{\infty} (1 - \mathcal{R}(\rho, \lambda_\gamma(0)))\, \mathrm{P}_{\|\boldsymbol{\Xi}\|}(d\rho) \geqslant 1 - \mathcal{R}(\tau', \lambda_\gamma 0) \tag{3.24}$$

It should be noted that, with the introduction of random variable ε, the associated notions of presence and absence probabilities (i.e. $\mathrm{P}[\varepsilon = 1]$ and $\mathrm{P}[\varepsilon = 0]$, respectively) seem to require prior information on the probability of the hypotheses, which is contrary to the standard Neyman-Pearson's principle. Actually, these quantities are unknown and play no important role in the constituted test. However, for the problem to be meaningful, $\mathrm{P}[\varepsilon = 1]$ must take its value in the interval $(0, 1)$. The introduction of random variable ε is simply for convenience in formulating the signal detection as an RDT problem.

In practice, it happens that the observation might not reduce to noise alone under the null hypothesis, but that there might still be presence of some unexpected signal that introduces a mismatch in observation modeling. This model mismatch might cause standard likelihood approach to violate the Neyman-Pearson's constraint on the false-alarm probability, but induces no such a problem in the RDT framework. To illustrate this aspect, the target signal is considered to be deterministic: $\boldsymbol{\Xi} = \boldsymbol{\xi}_1$ so that the Neyman-Pearson likelihood test applies.

To begin with, let us consider the ideal model in (3.22) with $\boldsymbol{\Xi} = \boldsymbol{\xi}_1$, $\|\boldsymbol{\xi}_1\| = \tau'$, i.e.:

$$\begin{cases} H_0: & \mathbf{Y} = \mathbf{X} \\ H_1: & \mathbf{Y} = \boldsymbol{\xi}_1 + \mathbf{X} \end{cases} \text{ with } \mathbf{X} \sim \mathcal{N}(0, \mathrm{C}), \|\boldsymbol{\xi}_1\| = \tau' \tag{3.25}$$

The standard Neyman-Pearson likelihood approach — namely test $\mathcal{T}_{\mathrm{NP}}$ — leads to compare the measure $\Lambda = \boldsymbol{\xi}_1^T \mathrm{C}^{-1}\mathbf{Y}$ to some threshold λ_{NP}, which is optimized in the sense that the ideal false-alarm probability given by $\mathcal{T}_{\mathrm{NP}}$ is: $\mathrm{P}^*_{\mathrm{fa}}[\mathcal{T}_{\mathrm{NP}}] \stackrel{\mathrm{def}}{=} \mathrm{P}\left[\boldsymbol{\xi}_1^T \mathrm{C}^{-1}\mathbf{X} > \lambda_{\mathrm{NP}}\right] = \gamma$. The corresponding detection probability is then:

$$\begin{aligned} \mathrm{P}^*_{\mathrm{d}}[\mathcal{T}_{\mathrm{NP}}] &\stackrel{\mathrm{def}}{=} \mathrm{P}\left[\boldsymbol{\xi}_1^T \mathrm{C}^{-1}(\boldsymbol{\xi}_1 + \mathbf{X}) > \lambda_{\mathrm{NP}}\right] \\ &= \mathrm{P}\left[\boldsymbol{\xi}_1^T \mathrm{C}^{-1}\mathbf{X} > (\lambda_{\mathrm{NP}} - \boldsymbol{\xi}_1^T \mathrm{C}^{-1}\boldsymbol{\xi}_1)\right] = 1 - \Phi\left(\frac{\lambda_{\mathrm{NP}} - \boldsymbol{\xi}_1^T \mathrm{C}^{-1}\boldsymbol{\xi}_1}{\sqrt{\boldsymbol{\xi}_1^T \mathrm{C}^{-1}\boldsymbol{\xi}_1}}\right) \end{aligned} \tag{3.26}$$

where $\Phi(.)$ is the cumulative distribution function (c.d.f.) of any standard normally distributed random variable.

Assuming now that in absence of signal $\boldsymbol{\xi}_1$, there are unavoidable fluctuations that cause the observation to randomly distort from zero, regardless whether additive noise $\mathbf{X}$ is present or not. Such a distortion, denoted by $\boldsymbol{\Xi}_0$, is random with unknown distribution in practice. However, its norm $\|\boldsymbol{\Xi}_0\|$ is generally bounded (almost surely) by

some positive value τ. The problem is actually:

$$\begin{cases} H_0: & \mathbf{Y} = \boldsymbol{\Xi}_0 + \mathbf{X} \\ H_1: & \mathbf{Y} = \boldsymbol{\xi}_1 + \mathbf{X} \end{cases} \text{ with } \mathbf{X} \sim \mathcal{N}(0, \mathrm{C}), \|\boldsymbol{\xi}_1\| = \tau', \|\boldsymbol{\Xi}_0\| \leqslant \tau \text{ (almost surely)} \tag{3.27}$$

If the likelihood test $\mathcal{T}_{\mathrm{NP}}$ is still applied, it might violate the Neyman-Pearson's constraint on the false-alarm probability. In fact, it can easily be shown that, as long as the random variable $\boldsymbol{\xi}_1^T \mathrm{C}^{-1} \boldsymbol{\Xi}_0$ is symmetrically distributed, the aforementioned constraint is always violated, i.e. $\mathrm{P}_{\mathrm{fa}}[\mathcal{T}_{\mathrm{NP}}] > \gamma$, for any $\gamma < 0.5$ (cf. the proof in Appendix A for details). On the contrary, with the same notation as before, by setting $\boldsymbol{\Theta} = \varepsilon\boldsymbol{\xi}_1 + (1-\varepsilon)\boldsymbol{\Xi}_0$, the detection problem of Eq. (3.27) is RDT. The test to be used is therefore $\mathcal{T}_{\lambda_\gamma(\tau)}$, which provides the probability of false-alarm:

$$\mathrm{P}_{\mathrm{fa}}\left[\mathcal{T}_{\lambda_\gamma(\tau)}\right] \stackrel{\mathrm{def}}{=} \mathrm{P}\left[\|\boldsymbol{\Xi}_0 + \mathbf{X}\| > \lambda_\gamma(\tau)\right] \leqslant 1 - \mathcal{R}(\tau, \lambda_\gamma(\tau)) = \gamma \tag{3.28}$$

The Neyman-Pearson's constraint on the false-alarm probability is thus respected in the RDT framework.

To illustrate these aspects, some numerical simulations with $d = 2$ and $\mathrm{C} = \sigma^2 \mathbb{1}_d$ were carried out. In these simulations, the unexpected distortion $\boldsymbol{\Xi}_0$ was randomly generated with normal distribution $\mathcal{N}(0, \sigma_0^2 \mathbb{1}_d)$ and τ was set to $\tau = 2\sigma_0$, which means that $\mathrm{P}[\|\boldsymbol{\Xi}_0\| < \tau] = 86.47\%$. Both the Neyman-Pearson likelihood test $\mathcal{T}_{\mathrm{NP}}$ and the RDT thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ were employed with different values of Signal-to-Noise Ratio (SNR) $\frac{\tau'}{\sigma}$ (10dB, 15dB, 20dB). The Signal-to-maximum-Distortion Ratio $\frac{\tau'}{\tau}$ was also set to the similar value $\frac{\tau'}{\tau} = 5$ (≈ 14dB) — i.e. the ratio $\frac{\sigma_0}{\sigma}$ equals -10dB, -5dB, 0dB respectively — which seemingly implies that the distortion caused by unexpected fluctuations is of very small magnitude. By such setting, we intend to show that, despite small distortion, the Neyman-Pearson likelihood test will fail. The detection results are reported in Fig. 3.1. On the one hand, Fig. 3.1 confirms that, for any $\gamma < 0.5$, the Neyman Pearson likelihood test actually yields a false-alarm rate higher than expected. On the other hand, although there is some loss in detection rate due to the unavoidable trade-off between the false-alarm and the detection probability, the detection in the RDT framework guarantees a false-alarm rate lower than the specified level γ and optimality of the detection with respect to the invariance of the problem.

It is also worth mentioning that, in case $\boldsymbol{\Xi}$ is not a deterministic signal $\boldsymbol{\xi}_1$ such that $\|\boldsymbol{\xi}_1\| = \tau'$ but a random one $\boldsymbol{\Xi}_1$ with unknown distribution and $\|\boldsymbol{\Xi}_1\| \geqslant \tau'$, the detection probability is lower bounded by $1 - \mathcal{R}(\tau', \lambda_\gamma(\tau))$. In other words, with the same simulation setting as before and $\boldsymbol{\Xi}_1$ randomly generated with any distribution that satisfies $\mathrm{P}[\|\boldsymbol{\Xi}_1\| \geqslant \tau'] = 1$, the resulting detection rate curves yielded by the thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ must be above those given in Fig. 3.1.

Moreover, in the above assessment, we have set up numerical simulations in such a way that $\mathrm{P}[\|\boldsymbol{\Xi}_0\| \leqslant \tau] = 86.47\%$, which implies we are far from assumption $\|\boldsymbol{\Xi}_0\| \leqslant$

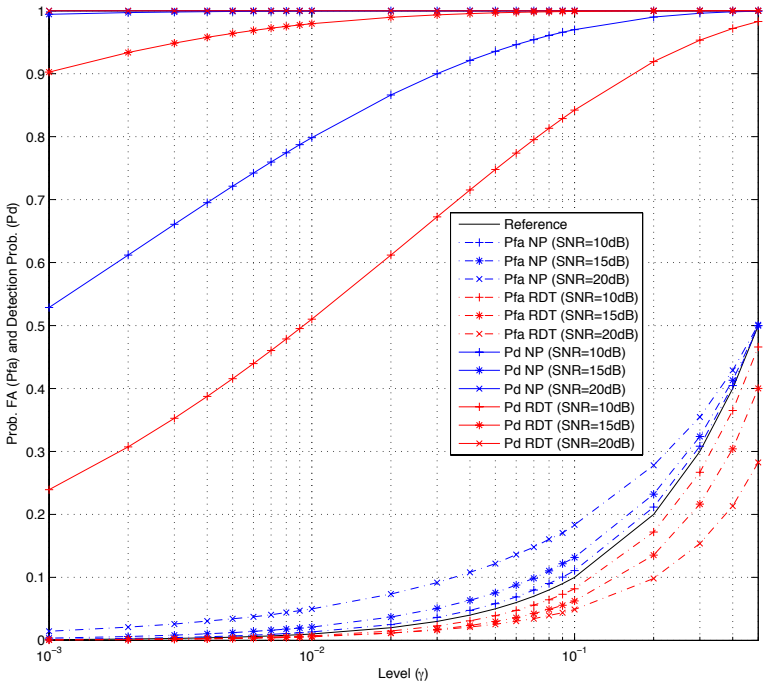


Figure 3.1 — Detection performance yielded by the Neyman-Pearson likelihood test $\mathcal{T}_{\mathrm{NP}}$ and the RDT thresholding test $\mathcal{T}_{\lambda_\gamma(\tau)}$ with different values of Signal-to-Noise Ratio (SNR) $\frac{\tau'}{\sigma}$ (10dB, 15dB, 20dB). The Signal-to-maximum-Distortion Ratio $\frac{\tau'}{\tau}$ was set to 5 ($\approx$ 14dB). The reference curve represents the Neyman-Pearson's constraint on the false-alarm probability.

τ (almost surely). Even in this case, RDT shows its advantage over the classical detection framework. Therefore, it is expected that the condition $\|\boldsymbol{\Xi}_0\| \leqslant \tau$ (almost surely) can significantly be relaxed in practice.

CHAPTER 4

Detection of signal deviation/distortion using RDT

Being motivated by various technical aspects in the mechanical ventilation monitoring where each observation channel is a one-dimensional temporal signal, we hereafter limit ourself to the one-dimensional case. The general case with multi-dimensional observations is postponed to future work, in which certain results given here for one-dimensional signal should be generalized. To begin with, let us consider the observation of the target temporal signal in noise. The signal is supposed to have undergone some deviation/distortion with respect to a nominal model. We are interested in verifying whether a significant deviation/distortion is present. Depending on the nature of the application, three problems are investigated:

Problem 1 [Dev.] - *Detection of signal deviations at specific instants:* Although the signal is observed for a long period of time, only the possible deviations at some specific instants — also called *critical instants* — are of interest. This type of detection is usually found in monitoring applications with scheduled sampling, testing and reaction. The detection of dynamic hyperinflation (i.e. AutoPEEP) during mechanical ventilatory support is among typical examples, in which the phenomenon can only be found at the end of expiratory phase of the respiratory signal.

Problem 2 [Chg.] - *Change point detection:* In this category of detection problems, changes present in the observed signal resulting from transition in hidden state of the signal source are considered. The detection of different waves in ECG (electrocardiography) signal, the detection of phase change in respiratory signal (flow, pressure, air volume) during assisted ventilation are examples amongst others.

Problem 3 [Dis.] - *Detection of signal distortion in a time interval:* Any significant deformation/distortion of a signal observed in a specific duration of time with

respect to a nominal model — i.e. a reference signal — is investigated. The monitoring with predefined waveform and a specific tolerance is typical application. The surveillance of carbon-dioxide level (P_{CO_2}) and the detection of asynchrony during mechanical ventilation are practical examples.

The aforementioned problems are classic and can be found in other works in the state of the art (for example, the change point detection in [Lai, 1995], the abrupt change detection in [Bassevile and Nikiforov, 1993, Lai and Shan, 1999, Fillatre, 2011], etc). In this work, these classical problems have been reformulated and cast into the RDT framework with very little prior information on observation supposed to be known. The resulting detection is robust and also provides statistical optimality.

4.1 Detection of signal deviations at specific instants - Extension of RDT in sequential detection framework

Using the same observation as before, the random process is supposed to be observed in additive gausian noise:

$$Y(t) = \Theta(t) + X(t) \tag{4.1}$$

where $Y(t)$ is the signal observation in noise $X(t)$. The distribution of random signal $\Theta(t)$ is generally unknown. In many practical applications, it is of much interest to know whether the clean (un-noised) random signal $\Theta(t)$ presents or not a significant deviation from some reference $f(t)$ at a specific critical instant t_c with respect to some tolerance τ, i.e.

$$\text{testing} \quad |\Theta(t_c) - f(t_c)| \leqslant \tau \quad \text{against} \quad |\Theta(t_c) - f(t_c)| > \tau \tag{4.2}$$

on the basis of observation $Y(t)$. This is an RDT problem in the sense given in Section 3.2.2 with $d = 1$.

4.1.1 Detection at one single critical instant

Although the detection could be carried out based exclusively on one sample of observation at critical instant t_c, taking its neighbor samples into account might improve the detection performance due to possible correlation among corresponding signal samples around the considered instant. To this end, let us consider $2L+1$ samples in the neighborhood of the sample at t_c, i.e. $Y(t_c - LT_s), ..., Y(t_c - T_s), Y(t_c), Y(t_c + T_s), ..., Y(t_c + LT_s)$ where T_s is the sampling period. It should be noted that, we consider here a neighborhood centered at t_c for the readiness sake; however, in practice, this

neighborhood could be any set of $2L+1$ samples, preferably consecutive, containing $Y(t_c)$. We then have the observation model in vector form as follows:

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{X} \tag{4.3}$$

where

$$\mathbf{Y} = \begin{bmatrix} Y(t_c - LT_s) \\ ... \\ Y(t_c - T_s) \\ Y(t_c) \\ Y(t_c + T_s) \\ ... \\ Y(t_c + LT_s) \end{bmatrix}, \ \mathbf{\Theta} = \begin{bmatrix} \Theta(t_c - LT_s) \\ ... \\ \Theta(t_c - T_s) \\ \Theta(t_c) \\ \Theta(t_c + T_s) \\ ... \\ \Theta(t_c + LT_s) \end{bmatrix} \text{ and } \mathbf{X} = \begin{bmatrix} X(t_c - LT_s) \\ ... \\ X(t_c - T_s) \\ X(t_c) \\ X(t_c + T_s) \\ ... \\ X(t_c + LT_s) \end{bmatrix}$$

are respectively the observation, the clean signal and the noise vectors. To profit from correlation among samples in the neighborhood of considered critical instant t_c to provide a better decision on $\Theta(t_c)$, we introduce vector $\mathbf{p} = [p_{-L}, ..., p_{-1}, p_0, p_1, ..., p_L]^T$ in such a way that vector $\mathbf{\Theta}$ can be factorized as:

$$\mathbf{\Theta} = \mathbf{p}\Theta(t_c) \tag{4.4}$$

This vector $\mathbf{p}$ actually represents the local waveform of signal in the neighborhood of the critical instant of interest. It should be noted that $p_0 = 1$. This constraint implies the scale invariance of the so-called local waveform vector. In many real-world applications, $\mathbf{p}$ is priorly known. The surveillance of carbon-dioxide level (P_{CO_2}) in the air mixture during ventilatory support is an example. Since P_{CO_2} is expected to remain constant and equal to a fixed value specified by clinician, it can be found that waveform vector $\mathbf{p}$ must equal to $[1, 1..., 1]^T$, which implies that, without any distortion and noise, all the samples of observation in the time interval of interest must be the same. For other applications, in which $\mathbf{p}$ is unknown but there is some known tendency of signal around the instant of interest, this waveform vector $\mathbf{p}$ can be estimated from the available samples of observations. Later on, in the next Chapter where the detection of AutoPEEP during mechanical ventilation is introduced, a method to estimate waveform vector $\mathbf{p}$ by regression from the available data will be shown.

By projecting observation vector $\mathbf{Y}$ onto the direction induced by waveform vector $\mathbf{p}$, we thus have:

$$\frac{\mathbf{p}^T\mathbf{Y}}{\|\mathbf{p}\|_2^2} = \Theta(t_c) + \frac{\mathbf{p}^T\mathbf{X}}{\|\mathbf{p}\|_2^2} \tag{4.5}$$

where $\|.\|_2$ is the standard Euclidean norm. Set $U = \frac{\mathbf{p}^T\mathbf{Y}}{\|\mathbf{p}\|_2^2}$ as new observation and $W = \frac{\mathbf{p}^T\mathbf{X}}{\|\mathbf{p}\|_2^2}$ as new noise, the equation is simplified as:

$$U = \Theta(t_c) + W \tag{4.6}$$

It should be mentioned that, by such a transformation, the problem becomes one-dimensional and possible correlation among noise samples plays no important role. Only the noise standard deviation will be needed. The problem is the same as before — i.e. testing $|\Theta(t_c) - f(t_c)| \leqslant \tau$ against $|\Theta(t_c) - f(t_c)| > \tau$, except that the variance of noise W is smaller than that of the original noise $X(t_c)$. In particular, if $X(t_c - LT_s), ..., X(t_c - T_s), X(t_c), X(t_c + T_s), ..., X(t_c + LT_s)$ are independent and identically distributed (iid) with variance σ_X, it can be proved that:

$$\sigma_W = \frac{\sigma_X}{\|\mathbf{p}\|_2} \tag{4.7}$$

and thus $\sigma_W < \sigma_X$ since $\|\mathbf{p}\|_2 > 1$. Furthermore, in most cases where $X(t_c - LT_s)$, ..., $X(t_c - T_s)$, $X(t_c)$, $X(t_c + T_s)$, ..., $X(t_c + LT_s)$ are independent and identically distributed (iid) and when the number of sample $2L+1$ is large enough, noise W becomes gaussian or approximately gaussian regardless what actual distribution of the original noise X might be (cf. Appendix C). Therefore, the gaussian assumption on original noise X could be relaxed in practice.

With regard to the aforementioned results, given an expected maximum value γ for the false-alarm rate, the detection can be carried out as follows:

$$\begin{cases} \text{If } |U - f(t_c)| > \lambda^*_\gamma \text{ then a significant deviation at } t_c \text{ is found.} \\ \text{If } |U - f(t_c)| \leqslant \lambda^*_\gamma \text{ then no significant deviation is present.} \end{cases} \tag{4.8}$$

The threshold λ^*_γ can be calculated in such a way that the upper bound of the false-alarm rate is always respected. However, the question is what optimality a test with respect to such a threshold might satisfy. In this respect, the RDT provides an optimal threshold λ^*_γ that yields a *maximal constant conditional power* test. More specifically, when W is centered gaussian noise, the threshold is given as $\lambda^*_\gamma = \sigma_W \lambda_\gamma(\frac{\tau}{\sigma_W})$, in which $\lambda_\gamma(\rho)$ is the unique solution in η to the equation:

$$1 - [\Phi(\eta - \rho) - \Phi(-\eta - \rho)] = \gamma \tag{4.9}$$

where $\Phi(.)$ is the cdf of any standard normal distributed random variable. It may also be noted here that, in this case, the map $\mathcal{R}$ is given by $\mathcal{R}(\rho, \eta) = \Phi(\eta - \rho) - \Phi(-\eta - \rho)$.

4.1.2 Repeated detections at multiple critical instants with extension of RDT in sequential detection framework

In many applications, it is required to repeatedly perform the detection at multiple critical instants $t_k, k \in \mathbb{N}$. The detection of AutoPEEP during mechanical ventilation, which investigates the deviation of flow signal at the end of expiratory phase of each breath, is a typical instance of such problem (see Section 5.2 for details). These critical

instants t_k might be regularly distributed in time, as in the AutoPEEP detection problem, or not. When the possible deviations at instants t_k are uncorrelated, the detection for each critical instants can be performed independently using the same proceeding as that described above. Otherwise, this correlation can be taken into account to provide a procedure with better detection performance. An extension of RDT in a sequential detection framework is proposed to deal with such a problem.

Let us consider multiple critical instants $t_k, k \geqslant 1$ for each of which a deviation detection must be carried out. In this application, the reference values $f(t_k)$ for the signal at instants t_k are assumed to be the same, i.e. $f(t_k) = f_0, \forall k$. Using the same aggregation scheme as in equation 4.5 for each critical instant t_k, we have:

$$U_k = \Theta(t_k) + W_k \quad \text{for} \quad k \geqslant 1 \tag{4.10}$$

It should be noted that, by using directly RDT for single detections at critical instants t_k, the false-alarm rate is upper-bounded by a specific value γ. It can also be shown that the detection probability is lower-bounded by γ. However, since γ is usually small, this lower-bound for the detection probability is of poor interest. It is thus required to improve the detection probability while still upper-bounding the false-alarm rate. To this end, two elements are taken into account.

First, two thresholds are introduced: one is calculated to restrict the false-alarm rate as usual, whereas the other is obtained by swapping the two hypotheses to limit the miss-detection rate. More specifically, we propose in this work to take both the problem *from-above*, i.e.

$$\begin{cases} h_0 : & |\Theta(t_k) - f_0| \leqslant \tau \\ h_1 : & |\Theta(t_k) - f_0| > \tau \end{cases}, \tag{4.11}$$

and the problem *from-below*, i.e.

$$\begin{cases} h'_0 : & |\Theta(t_k) - f_0| > \tau \\ h'_1 : & |\Theta(t_k) - f_0| \leqslant \tau \end{cases} \tag{4.12}$$

into account. As suggested by the RDT framework, on the one hand, the thresholding test $\mathcal{T}_{\lambda^*_\gamma}$ from-above the optimal threshold $\lambda^*_\gamma = \sigma_W \lambda_\gamma(\tau/\sigma_W)$ is defined by:

$$\mathcal{T}_{\lambda^*_\gamma}(U_k) = \begin{cases} 1 & \text{if} \quad |U_k - f_0| > \lambda^*_\gamma \\ 0 & \text{if} \quad |U_k - f_0| \leqslant \lambda^*_\gamma \end{cases} \tag{4.13}$$

and has size γ for the problem from-above (4.11) so that:

$$\mathrm{P}\left[|U_k - f_0| > \lambda^*_\gamma\right] \leqslant \gamma \text{ when } |\Theta(t_k) - f_0| \leqslant \tau; \tag{4.14}$$

on the other hand, the thresholding test $\mathcal{T}'_{\lambda^*_{1-\gamma}}$ from-below the optimal threshold $\lambda^*_{1-\gamma} = \sigma_W \lambda_{1-\gamma}(\tau/\sigma_W)$ is given by:

$$\mathcal{T}'_{\lambda^*_{1-\gamma}}(U_k) = \begin{cases} 1 & \text{if} \quad |U_k - f_0| \leqslant \lambda^*_{1-\gamma} \\ 0 & \text{if} \quad |U_k - f_0| > \lambda^*_{1-\gamma} \end{cases} \tag{4.15}$$

and has size γ for the problem from-below (4.12), i.e.:

$$\mathrm{P}\left[|U_k - f_0| < \lambda^*_{1-\gamma}\right] \leqslant \gamma \text{ when } |\Theta(t_k) - f_0| > \tau \tag{4.16}$$

For $\gamma < 0.5$, it can be shown that $\lambda^*_{1-\gamma} < \lambda^*_\gamma$ (cf. Appendix B). It is also worth mentioning that γ is usually set to be small. In practice, $\gamma = 0.001, 0.005, 0.01, 0.05$ are typical values. To combine the two constraints on false-alarm rate given by the two aforementioned tests, a new thresholding test with dual thresholds $\lambda^*_{1-\gamma}$ and λ^*_γ is formulated as follows:

$$\mathcal{T}''_{[\lambda^*_{1-\gamma},\lambda^*_\gamma]}(U_k) = \begin{cases} 1 & \text{if} \quad |U_k - f_0| > \lambda^*_\gamma \\ 0 & \text{if} \quad |U_k - f_0| \leqslant \lambda^*_{1-\gamma} \\ ? \text{ (not decided yet)} & \text{otherwise} \end{cases} \tag{4.17}$$

for testing the null hypothesis $[h_0 : |\Theta(t_k) - f_0| \leqslant \tau]$ against the alternative one $[h_0 : |\Theta(t_k) - f_0| > \tau]$. It is directly inherited from (4.14) and (4.16) that, once the decision has been made by the dual-threshold test $\mathcal{T}''_{[\lambda^*_{1-\gamma},\lambda^*_\gamma]}$ proposed in (4.17), the probability of false-alarm ($\mathrm{P_{fa}}$) is limited to be lower than the specified value γ and the detection probability is guaranteed to be higher than $1-\gamma$, i.e:

$$\begin{array}{rcl} \mathrm{P_{fa}}\left[\mathcal{T}''_{[\lambda^*_{1-\gamma},\lambda^*_\gamma]}\right] & \leqslant & \gamma \\ \mathrm{P_{d}}\left[\mathcal{T}''_{[\lambda^*_{1-\gamma},\lambda^*_\gamma]}\right] & \geqslant & 1-\gamma \end{array} \tag{4.18}$$

Furthermore, statistical optimalities provided by $\mathcal{T}_{\lambda^*_\gamma}$ and $\mathcal{T}'_{\lambda^*_{1-\gamma}}$ are also inherited by $\mathcal{T}''_{[\lambda^*_{1-\gamma},\lambda^*_\gamma]}$.

Second, when it turns out that the proposed dual-threshold test is unable to provide a decision with enough confidence based on a single observation U_k — in other words, $\lambda^*_{1-\gamma} < |U_k - f_0| \leqslant \lambda^*_\gamma$ —, the decision is then postponed until more observations are acquired. A mechanism to aggregate these multiple consecutive observations into a decision is then required. Such mechanism must satisfies that, the more observations are collected, the more information that supports a more confident decision can be inferred. Let $U_1, U_2, ..., U_K$ be K consecutive observations available for the decision. One approach is to average over these observations, as follows:

$$U_{1:K} = \Theta_{1:K} + W_{1:K} \tag{4.19}$$

where:

$$U_{1:K} = \tfrac{1}{K}\sum_{k=1}^{K} U_k \quad , \quad \Theta_{1:K} = \tfrac{1}{K}\sum_{k=1}^{K} \Theta(t_k) \quad \text{and} \quad W_{1:K} = \tfrac{1}{K}\sum_{k=1}^{K} W_k.$$

It is worth mentioning that, as long as $W_k \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_W)$, we have $W_{1:K} \sim \mathcal{N}(0, \sigma^2_{W,K})$ and that $\sigma_{W,K} = \frac{\sigma_W}{\sqrt{K}}$ is strictly decreasing with respect to the number K of observations to be aggregated.

Assuming that the clean signal does not vary too much for K instants $t_k, k = 1..K$, the deviation detection can be seen as:

$$\text{testing} \quad |\Theta_{1:K} - f_0| \leqslant \tau \quad \text{against} \quad |\Theta_{1:K} - f_0| > \tau, \tag{4.20}$$

The proposed dual-threshold test is then performed on the basis of $U_{1:K}$. In particular, given a desired maximum value γ for the false-alarm rate, the detection is attempted with decision rule $[d(U_{1:K})]$ defined by:

$$[d(U_{1:K})] \begin{cases} \textbf{If} & |U_{1:K} - f_0| > \lambda^{(h)}_{1:K} & \textbf{then} & \text{a significant deviation is found.} \\ \textbf{If} & |U_{1:K} - f_0| \leqslant \lambda^{(\ell)}_{1:K} & \textbf{then} & \text{no significant deviation is present.} \\ \textbf{If} & \lambda^{(\ell)}_{1:K} < |U_{1:K} - f_0| \leqslant \lambda^{(h)}_{1:K} & \textbf{then} & \text{No decision is made yet. Decision} \\ & & & \text{is delayed until next critical instant} \end{cases} \tag{4.21}$$

The two thresholds are given by:

$$\begin{aligned} \lambda^{(h)}_{1:K} &= \sigma_{W,K} \lambda_{\gamma} \left(\tfrac{\tau}{\sigma_{W,K}} \right) \\ \lambda^{(\ell)}_{1:K} &= \sigma_{W,K} \lambda_{1-\gamma} \left(\tfrac{\tau}{\sigma_{W,K}} \right) \end{aligned} \tag{4.22}$$

in which $\lambda_{\gamma}(\rho)$ (resp. $\lambda_{1-\gamma}(\rho)$) is the unique solution in η to the equation $1-\mathcal{R}(\rho,\eta) = \gamma$ (resp. $1 - \mathcal{R}(\rho,\eta) = 1 - \gamma$), where $\mathcal{R}(\rho,\eta) = \Phi(\eta - \rho) - \Phi(-\eta - \rho)$ and $\Phi(.)$ is the cumulative distribution function (cdf) of any standard normally distributed random variable.

With respect to these two aforementioned elements, in a sequential decision framework, the detection is carried out as follows. At the beginning, decision rule $[d(U_{1:1})]$ is used as the first attempt to provide decision based solely on the observation for the first critical instant t_1. If a decision can be made with enough confidence (i.e. either $|U_{1:1} - f_0| > \lambda^{(h)}_{1:1}$ or $|U_{1:1} - f_0| \leqslant \lambda^{(h)}_{1:1}$), the process stops with that decision. Otherwise, if the decision cannot be made yet (i.e. $\lambda^{(\ell)}_{1:1} < |U_{1:1} - f_0| \leqslant \lambda^{(h)}_{1:1}$), it will be delayed until the next observation (i.e. U_2) is obtained and the test is performed based on $U_{1:2}$ using decision rule $[d(U_{1:2})]$. If the decision still cannot be accomplished, it will be delayed again until the next observation, where the decision rule $[d(U_{1:3})]$ is used and so on. The process is iterated until the decision is made. Then the process is restarted for a new sequence of observations.

It can be seen that $\lambda^{(\ell)}_{1:K} < \lambda^{(h)}_{1:K}$ for any $0 < \gamma < 0.5$ (cf. Appendix B). As shown in Figure 4.1, the two thresholds tend to tolerance τ when $\sigma_{W_{1:K}}$ tends to 0 (cf. Appendix B for the proof). It should also be noted that $\sigma_{W,K} = \frac{\sigma_W}{\sqrt{K}}$ decreases with respect to K and tends to 0 when K tends to $+\infty$. Therefore, the decision will probably be made after a finite number of observations is acquired. However this finite number is not known prior to the decision to be made and it may be too high to admit the assumption that the signal of interest does not vary too much within the considered K consecutive

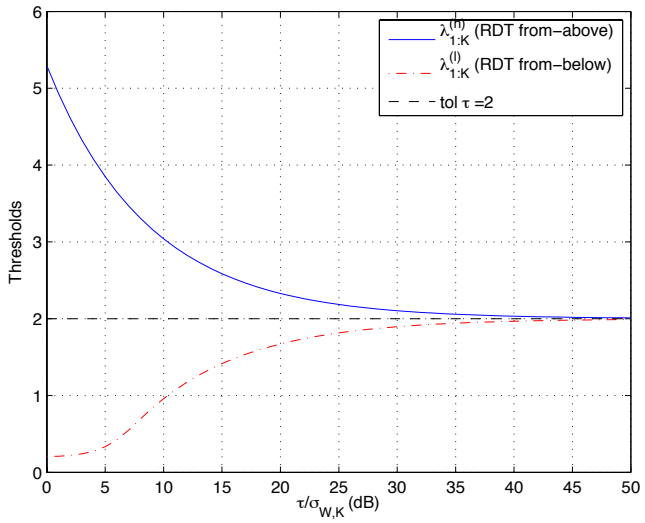


Figure 4.1 — Thresholds convergence. This figure illustrates the convergence of the two thresholds in Sequential RDT framework. This convergence suggests that, in sequential RDT framework, the decision will probably be made after a finite number of samples are acquired.

observations. Moreover, a high number of observations to be acquired may yield an unacceptable delay-to-decision. One simple solution is to limit the number of critical instants to some value M. If M observations for the detection at M critical instants have been acquired but no decision has been made, a *hard* decision is performed. To privilege the upper-bounding of the false-alarm rate, threshold $\lambda^{(h)}_{1:M}$ is used. Once it is required, the hard decision is carried out by:

$$[d_h(U_{1:M})] \begin{cases} \textbf{If} \quad |U_{1:M} - f_0| > \lambda^{(h)}_{1:M} \quad \textbf{then} \quad \text{a significant deviation is found.} \\ \textbf{If} \quad |U_{1:M} - f_0| \leqslant \lambda^{(h)}_{1:M} \quad \textbf{then} \quad \text{no significant deviation is present.} \end{cases} \tag{4.23}$$

The value M must be chosen so that the aforementioned assumption holds and the delay-to-decision is still in an acceptable range.

It could also be remarked that, in this section, the same level γ is used for both thresholding tests from-above and from-below. In fact, this is not necessary. Two different levels, one for each type of decision error, can be employed. However, the additional one — aimed at limiting the miss-detection rate and, therefore, lower-bounding the detection rate— is less meaningful because the sequential process might be stopped by a hard decision before this bound can be attained.

4.2 Change point detection

In practice, there are cases in which the observed signal needs to be segmented into different portions — called phases — whose nature depends on the corresponding

hidden state of the signal source of interest. Such segmentation task requires change points yielded by internal hidden state transition to be detected. The detection of phase changes in periodic signals — such as electrocardiography (ECG) signal, respiratory signal, etc — represents a class of examples. In this section, let us take the flow signal during assisted mechanical ventilation as an illustration (cf. Figure 4.2). In this example, the detection of change points from inspiratory phase to expiratory phase and vice versa is investigated.

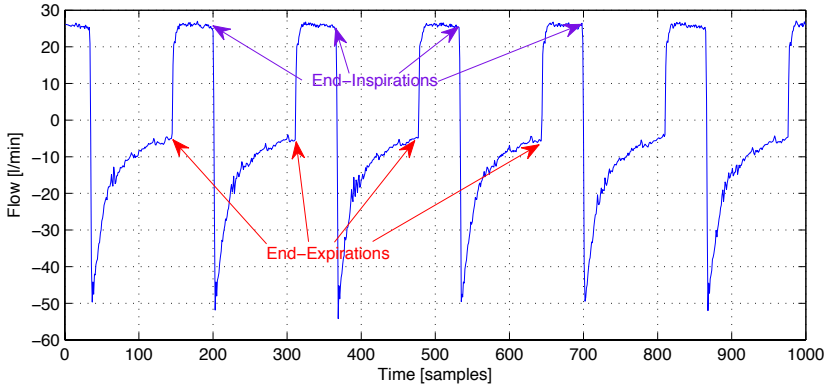


Figure 4.2 — Flow signal captured from patient undergoing mechanical ventilation with respiratory phase changes.

It could be noted that the signal in each phase can be assumed to be regular without any abrupt changes. Only transition in internal phases can cause irregularities in signal. The detection of phase changes then resorts to detecting such irregularities based on observation of signal in noise. In this section, we propose an approach to accomplish such task using RDT framework. The idea is to perform successively two steps: first, the observed signal is transformed into another domain in which the irregularities in the signal can be more emphasized and, second, the RDT thresholding test with optimal threshold is carried out to reveal significant abrupt changes in signal regularity, that correspond to phase changes. The tolerance and the level are carefully specified so as to avoid any effects yielded by unexpected noise.

Since the wavelet transform is a powerful processing tool to retrieve irregularities in a signal (cf. [Donoho and Johnstone, 1994] amongst others), it can be used to carry out the desired change detection. Let y_n be the observed signal. By using wavelet transformation, phase changes in signal result in significant peaks in the detail band (high band) signal of the wavelet coefficients. The change detection will be performed by thresholding these peaks in the detail bands of the wavelet transform coefficients. Figure 4.3 shows an example of these peaks for the detection of phase change in respiratory flow signal under ventilatory support. In this example, the discrete stationary wavelet transform was used and the number of wavelet decomposition levels was set to $K = 3$.

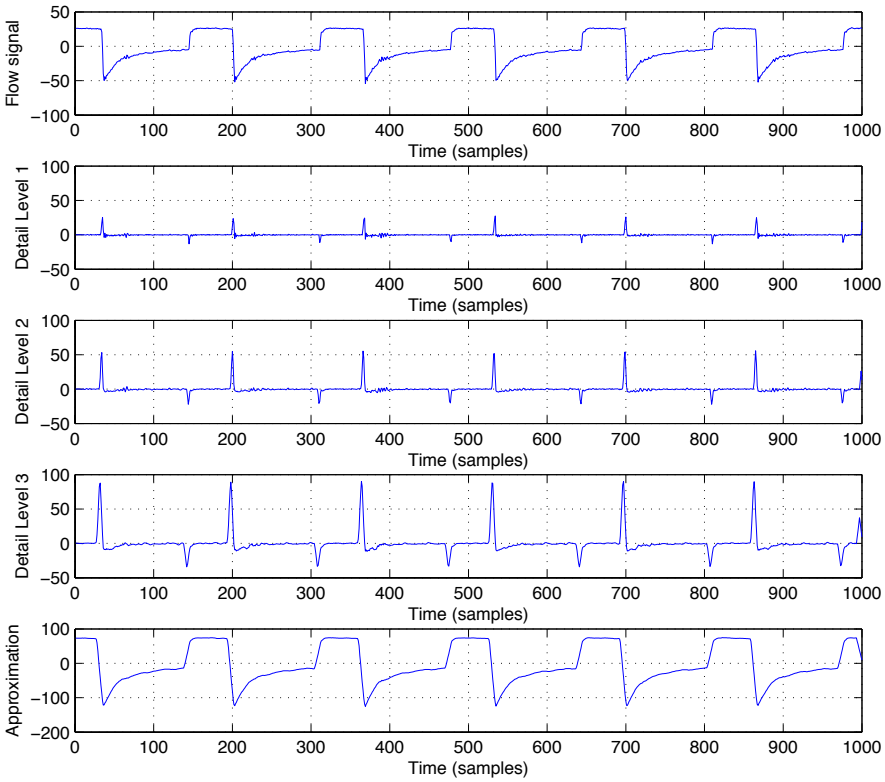


Figure 4.3 — Wavelet decomposition of the flow signal. The peaks in detail bands correspond to changes from inspiratory phase to respiratory phase and vice versa.

More specifically, let us consider a detail band, say level-2 detail band for instance. This detail band signal, namely y_D, is composed of noise and peaks, which represent the irregularities in the original signal. It has been shown that when the number of wavelet decomposition level is large enough, the noise in detail band tends to be gaussian and decorrelated [Atto and Pastor, 2010, Leporini and Pesquet, 1999, Pastor and Gay, 1995]. Therefore, each coefficient in the detail band can be modeled as

$$y_D = f_D + x_D \tag{4.24}$$

where f_D is a signal wavelet coefficient and x_D is gaussian noise. Let σ_D be the standard deviation of this noise and let N be the number of coefficients. It was shown in [Berman, 1992, Mallat, 1999, Serfling, 1980] that the maximum absolute value of noise when N tends to infinity is the universal threshold, defined by:

$$\lambda_u(N) = \sigma_D\sqrt{2\ln N}. \tag{4.25}$$

The universal threshold $\lambda_u(N)$ can also be thought of as the minimum absolute value of the signal (cf. [Pastor and Atto, 2010]). Therefore, the problem amounts to testing the peak absolute value with respect to $\lambda_u(N)$, that is:

$$\text{testing} \quad |y_D| \leqslant \lambda_u(N) \quad \text{against} \quad |y_D| > \lambda_u(N) \tag{4.26}$$

This is an RDT problem in the sense given in Section 3.2.2. The peaks in the detail band can thus be detected using the test:

$$\mathcal{T}_{\lambda_{\mathrm{RDT}}}(y_D) = \begin{cases} 1 & \text{if} \quad |y_D| > \lambda_{\mathrm{RDT}} \\ 0 & \text{if} \quad |y_D| \leqslant \lambda_{\mathrm{RDT}} \end{cases}$$

with threshold height $\lambda_{\mathrm{RDT}} = \sigma_D \lambda_\gamma \left(\frac{\lambda_u(N)}{\sigma_D} \right)$ where $\lambda_\gamma(\rho)$ is the unique solution in η to equation $1 - [\Phi(\eta - \rho) - \Phi(-\eta - \rho)] = \gamma$ and $\Phi(.)$ is, as before, the cdf of any standard normal distributed random variable.

It is worth mentioning that level γ can be used to control the sensitivity of the phase change detector. This level γ is usually set to be small in practice, for example $\gamma = 10^{-4}, 10^{-5}, 10^{-7}$, etc. As long as noise standard deviation σ_D is concerned, it can be estimated using any method in the state-of-the-art literature. For example, one could consider either the MAD (median absolute deviation) or the DATE (d-dimensional adaptive trimming estimator) mentioned in Section 5.2.4 below. Figure 4.4 gives a typical result of the phase change detection obtained by proceeding as described above.

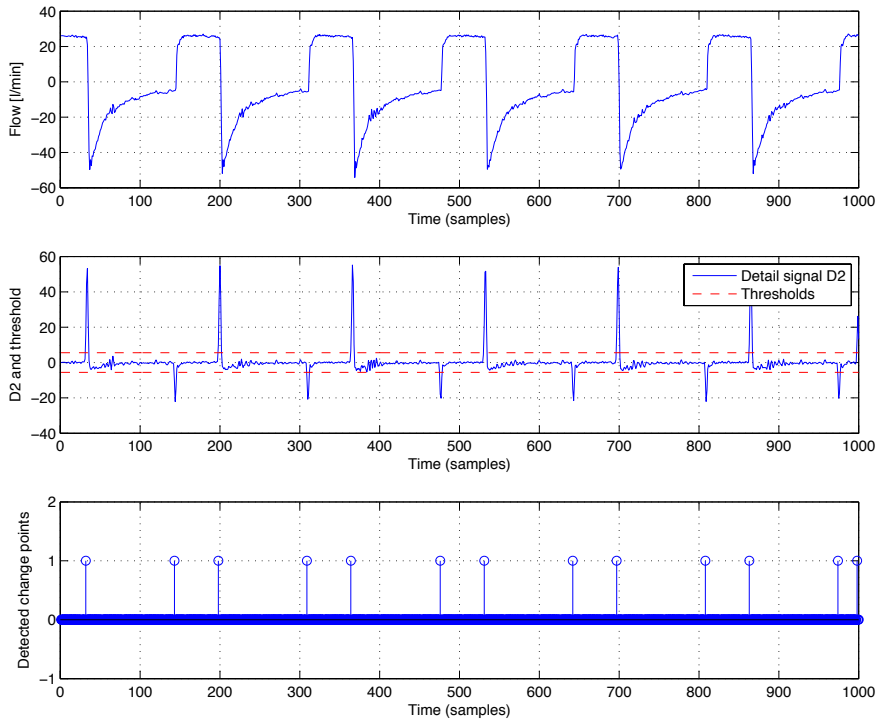


Figure 4.4 — Phase change detection using Wavelet transform, universal threshold and RDT

As aforementioned, when the number N of wavelet coefficients is large enough, the universal threshold $\lambda_u(N)$ can be considered as the maximum noise amplitude [Pastor and Atto, 2010]. Therefore, one might expect that $\lambda_u(N)$ can be directly used in a

thresholding test to filter out the noise and thus, the remaining significant coefficients corresponding to signal change points can be revealed. However, this direct approach can hardly provide good detection in practice since the false-alarm rate, which depends mainly on the probability that the noise coefficients amplitude is bounded by $\lambda_u(N)$, might be too high to be meaningful. Moreover, since the length of the observing window is finite and usually short in real-time applications, the number N of wavelet coefficients might be too limited and, consequently, the bound $\lambda_u(N)$ for noise amplitude might be severely violated. On the contrary, by using RDT, the condition on the bound value for noise amplitude can be significantly relaxed (cf. Section 3.3). Therefore, the resulting detector tolerates such limitations and provides good results.

4.3 Detection of signal distortion in a time interval

In this category of problem, the detection of any significant deformation/distortion of a signal with respect to a nominal model (i.e. a reference signal) during a certain interval of time is of interest. The detection of asynchrony during ventilatory support is an example of application. In Figure 4.5, a distortion of flow signal during expiratory-phase of the breath is present. This distortion is caused by an ineffective effort of patient in triggering the mechanical ventilator. In such case, the inspiration — periodically triggered by the ventilator itself on the basis of prior settings specified by clinician — will then start later than expected.

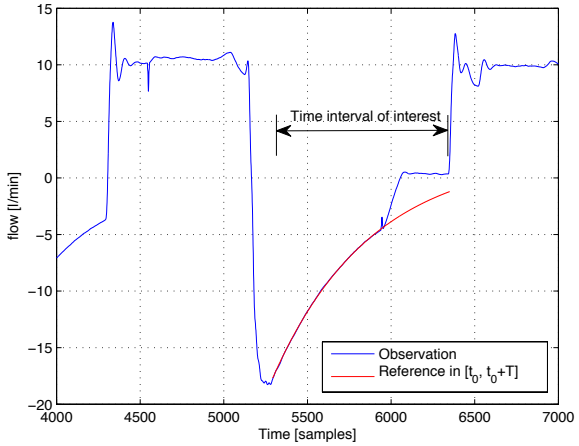


Figure 4.5 — Detection of asynchrony during ventilatory support. In this example, the signal distortion during the considered time interval represents an ineffective effort by patient undergoing mechanical ventilation. The presence of such ineffective effort entails that the triggering — automatically performed by the ventilator itself — happens later than expected.

To begin with, the same signal observation model is used: $Y(t) = \Theta(t) + X(t)$, where $Y(t)$ is the observation of signal $\Theta(t)$ in noise $X(t)$. The clean (unnoised) signal

$\Theta(t)$ is characterized by a known deterministic nominal model $f_0(t)$ suffering from an unknown random deformation or distortion $\Delta(t)$ whose distribution is also unknown:

$$\Theta(t) = f_0(t) + \Delta(t) \tag{4.27}$$

It is of interest to know whether or not, during a specific time interval $[t_0, t_0 + T]$, the clean signal $\Theta(t)$ presents or not a significant deformation/distortion from the nominal model $f_0(t)$ with respect to tolerance τ.

Let us considered L samples of the signal in the time interval of interest $[t_0, t_0 + T]$. The observation vector is $\mathbf{Y} = [Y_1, Y_2, ..., Y_L]^T$. With respect to the signal model, we have:

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{X} \tag{4.28}$$

where $\mathbf{\Theta} = [\Theta_1, \Theta_2, ..., \Theta_L]^T$ is the clean signal vector and $\mathbf{X}$ is the noise vector. We also have:

$$\mathbf{\Theta} = \mathbf{f}_0 + \mathbf{\Delta} \tag{4.29}$$

where $\mathbf{f}_0 = [f_{0,1}, f_{0,2}, ..., f_{0,L}]^T$ is known nominal (reference) signal vector and $\mathbf{\Delta} = [\Delta_1, \Delta_2, ..., \Delta_L]^T$ is the unknown deformation/distortion vector. The problem is then

$$\text{testing} \quad \|\mathbf{\Theta} - \mathbf{f}_0\| \leqslant \tau \quad \text{against} \quad \|\mathbf{\Theta} - \mathbf{f}_0\| > \tau \tag{4.30}$$

on the basis of observation $\mathbf{Y}$. Norm $\|.\|$ is chosen so as to compensate the variation induced by the noise covariance matrix. Typically, $\|.\|$ is the Malahanobis norm. This is directly an RDT problem. Given a desired maximum value γ for the false-alarm rate , the signal deformation/distortion detection is carried out as follows:

$$\begin{cases} \textbf{If} \quad \|\mathbf{Y} - \mathbf{f}_0\| > \lambda^*_\gamma \quad \textbf{then} & \text{a significant deformation/distortion} \\ & \text{in } [t_0, t_0 + T] \text{ is found.} \\ \textbf{If} \quad \|\mathbf{Y} - \mathbf{f}_0\| \leqslant \lambda^*_\gamma \quad \textbf{then} & \text{no significant deformation/distortion} \\ & \text{is present.} \end{cases} \tag{4.31}$$

where threshold λ^*_γ is also calculated in such a way that the upper bound of the false-alarm rate is always respected. More precisely, λ^*_γ is derived from the condition:

$$1 - F_{\|\Delta + \mathbf{X}\|}(\lambda^*_\gamma) \leqslant \gamma \quad \text{for all } \Delta \text{ that } \|\Delta\| \leqslant \gamma \tag{4.32}$$

where $F_{\|\Delta+\mathbf{X}\|}(.)$ is still the cumulative distribution function (cdf) of random variable $\|\Delta + \mathbf{X}\|$. As long as $\mathbf{X}$ is centered gaussian noise with covariance matrix $\mathbf{C}$, i.e. $\mathbf{X} \sim \mathcal{N}(0, \mathbf{C})$, the threshold is is given by $\lambda^*_\gamma = \lambda_\gamma(\tau)$, in which $\lambda_\gamma(\rho)$ is the unique solution in η to the equation:

$$1 - \mathcal{R}(\rho, \eta) = \gamma \tag{4.33}$$

where $\mathcal{R}(\rho, .)$ is the cumulative distribution function of the square root of any random variable that follows the non-central Chi-squared (χ^2) distribution with L degrees of freedom and non-central parameter ρ^2.

Reference $f_0(t)$ is supposed to be known prior to the detection. This reference is actually the desired waveform of the signal under consideration, i.e. it represents what the signal should be. The tolerance τ specifies to what extent the deviation of the observed signal from the reference is still acceptable. In many practical cases where the signal is periodic, $f_0(t)$ can be estimated by a complementary processing on the basis of the first periods that the system operator has verified and considered as the reference, i.e. distortion-free.

The three detection problems presented above are very general and can be found in many real-world applications. In the next chapter, these three problems will be found in the detection of AutoPEEP/Asynchrony during mechanical ventilation for an automatic monitoring framework. Other domains of application can also be investigated, including fault detection, structural health monitoring, remote sensing, robotics, etc.

CHAPTER 5

Application to mechanical ventilation system monitoring: AutoPEEP/Asynchrony detection

5.1 Introduction

The objective of mechanical (or artificial) ventilation is to assist or to completely replace the spontaneous breathing of the patient by a ventilator when the patient breathing becomes inefficient or, in some cases, absent. Mechanical ventilation is routinely used in emergency wards, operating rooms, or intensive care units. It is also used at home or in nursing/rehabilitation institutions, particularly for patients suffering from chronic illness and for those whose spontaneous breathing is insufficient. Unfortunately, imperfect interaction between patient and ventilator is very common. It has been shown that patient-ventilator mismatching is very frequently exhibited in both intubated patients receiving pressure support ventilation [Thille et al., 2006] and those undergoing non-invasive ventilation [Vignaux et al., 2009]. Among these abnormalities, dynamic hyperinflation and patient-ventilator asynchrony are very frequent, but are not yet detected in routine by currently used mechanical ventilators. Such imperfect interaction may generate incomplete ventilatory assistance, or even increased respiratory effort, thus generating deleterious adverse events and decreased prognosis. Therefore, the detection — possibly followed by an appropriate correction — of such abnormalities at patient-ventilator interface is necessary.

It has been demonstrated in literature that the graphical curves (flow, airway pressure and air volume) available on most recent mechanical ventilators provide much information to analyze the patient-ventilator interface [Roeseler et al., 2010]. By visually monitoring these curves, patient-ventilator mismatching can be observed and

detected by the clinician. Using the same inputs, automatic detection of ventilatory abnormalities can also be investigated. Various automatic detection algorithms, either embedded in a mechanical ventilator to detect ineffective triggering and double triggering [Mulqueeny et al., 2007], or recently in a computerized monitoring system (BetterCare) to determine ineffective respiratory efforts during expiration [Blanch et al., 2012], have been reported with positive results. However, to the best of our knowledge, the automatic detection of other types of ventilatory abnormalities, including dynamic hyperinflation, has not yet been adequately considered.

In this chapter, we address automatic detection of dynamic hyperinflation, a common ventilatory abnormality that usually occurs in patients with acute severe asthma or chronic obstructive pulmonary disease. This phenomenon is hereafter called AutoPEEP (Auto-Positive End Expiratory Pressure) with some slight language abuse for reading simplicity. The presence of AutoPEEP basically indicates an insufficient expiratory time. The amount of time given over to expiration therefore needs to be lengthened, either by reducing the respiration rate or by decreasing the inspiratory time, or both. AutoPEEP can be measured at the patient's bedside by using the pressure transducer of the ventilator. However, this quantification requires intervention from the therapist, who must perform an expiratory pause, in order to monitor tele-expiratory pressure [Blanch et al., 2005]. On the contrary, although not readily quantifiable, AutoPEEP can easily be recognized on the expiratory portion of the flow waveform. In particular, if expiratory flow does not return to zero before the next inspiration, AutoPEEP is present. This seemingly simple detection, however, requires the eye of an expert clinician at the patient's bedside. Using flow signal as the input, an automatic detection of AutoPEEP (dynamic hyperinflation) due to either expiratory flow limitation and/or inappropriate ventilatory cycling should be helpful to optimize care. Our focus is thus early detection of AutoPEEP for continuous monitoring of the patient-ventilator interface. AutoPEEP detection is performed by either RDT or its extension in sequential framework on the flow signal captured from the patient-ventilator interface. More specifically, it involves testing the norm of a signal observed in noisy condition with respect to a certain tolerance fixed by users on the basis of their know-how and/or experience of the domain (cf. Chapter 3). Two detectors, one based on each single breath and the other based on sequential decision on consecutive breaths are proposed. Practical aspects, including phase change detection and parameter estimation are considered as well. The performance assessment is provided in three levels. First, the detection performance of the proposed detectors will be illustrated with data synthesized on computer. Then, further evaluation is performed on data derived from a respiratory system analog which is arranged in a setting similar to that in practice. Finally, an *ex-vivo* performance assessment on retrospective data acquired from patients is carried out. *In-vivo* analysis with possible closed-loop tests will be postponed to a future work.

5.2 Automatic detection of AutoPEEP

AutoPEEP can be visually observed and detected through flow signal. Figure 5.1 shows an example of flow signal with AutoPEEP captured during mechanical ventilation on a patient.

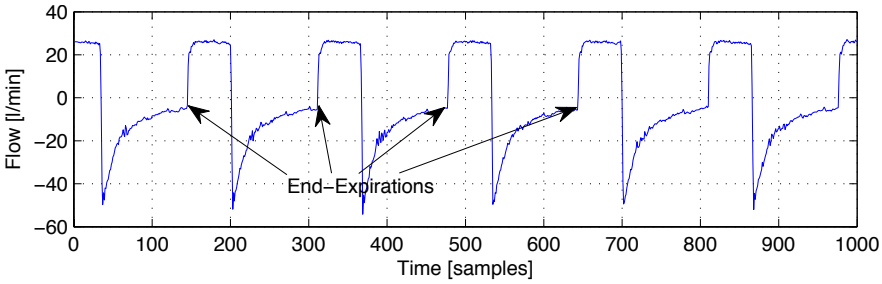


Figure 5.1 — An example of flow signal. This signal was recorded during the assisted mechanical ventilation on a patient. The (blue) curve shows a typical waveform of flow signal with squared inspiratory phase. The arrows point to some end-expiration instants where the markers for AutoPEEP detection are present.

Let f_t be the clean flow signal. AutoPEEP can be regarded as the non-return of the flow signal at the end of each expiratory phase to the null value. In practice, during the observation of the air flow, various factors might get involved, including the mechanical vibration of the air tube, the patient movement, the electro-magnetic interference, etc. Therefore, the flow signal at the end of the expiratory phase will never be exactly zero, even in absence of noise. Testing directly the hypothesis $f_{t_k} = 0$ against $f_{t_k} \neq 0$, where t_k is the end-expiration instant of the considered breath, might thus not be realistic. A tolerance $\tau > 0$ is then introduced to take into account possible distortions on the signal under consideration. Given τ, the problem is then the testing of $|f_{t_k}| \leqslant \tau$ versus $|f_{t_k}| > \tau$ based on the flow signal observation in presence of noise. This tolerance τ is specified by the clinician. Its value is usually derived from his/her expertise of the domain. Other technical factors could also be taken into account, such as: the flow sensor precision, the dynamic range of the signal, etc. Multiple values of τ could also be employed to provide a semi-quantitative evaluation of persisted AutoPEEP on patient.

5.2.1 System overview

With respect to the discussion above, a platform for automatic detection of AutoPEEP based on a noisy observation of the flow signal can be developed. Figure 5.2 depicts such a platform. The main processing components include: the data acquisition and Serial/Parallel conversion, the phase-change detector, the estimators and the AutoPEEP detector. These components are briefly presented as follows before being detailed in the sequel.

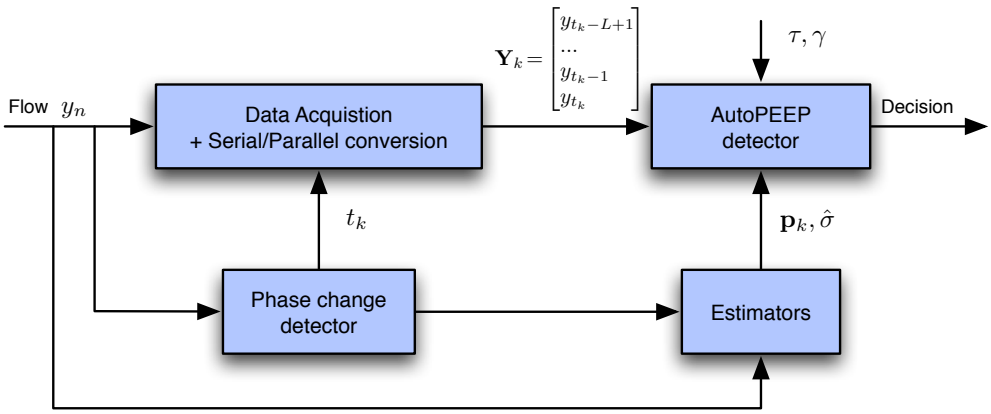


Figure 5.2 — The Automatic AutoPEEP Detection platform - System overview. The platform functions on the basis of respiratory flow signal. For each end-expiration t_k identified by the Phase change detector, L end-expiratory flow samples are logged to form an observation vector. Based on observations $\mathbf{Y}_k$ provided by the Data acquisition/conversion module and parameters $\mathbf{p}_k, \hat{\sigma}$ given by the Estimators, the AutoPEEP Detector performs an optimal testing with respect to specified tolerance τ and level γ to decide whether or not an AutoPEEP is present.

Data acquisition and Serial/Parallel conversion This very-first module acquires the discrete flow signal y_n provided by the ventilator or by an independent flow sensor installed inside the air-tube during the mechanical ventilation. Although every flow datum is acquired, only end-expiration flow data of each breath is useful for the detection of AutoPEEP. When the end-expiration instant t_k of the k-th breath is provided by the phase change detector, the Data Acquisition and Serial/Parallel conversion module will log L samples at the end of the expiratory phase to form the observation vector $\mathbf{Y}_k = [y_{t_k-L+1}, ..., y_{t_k-1}, y_{t_k}]^T$ for the k-th breath. This output observation vector $\mathbf{Y}_k$ is finally injected into the AutoPEEP detector module.

Respiration phase change detection The main role of this module is to detect the end-expiration of each breath and provide this instant to trigger the data logging process and the Serial/Parallel conversion described above. This can also be regarded as a breath detector, which separates the continuous flow signal into different breaths.

Estimators This module consists of two estimators, which estimate necessary parameters for the AutoPEEP detection algorithms. These parameters are the so-called waveform vector ($\mathbf{p_k}$ for the k-th breath) and the noise standard deviation estimate ($\hat{\sigma}$). The waveform vector will be used to aggregate multi-samples at the end of the expiratory phase of a breath into a decision (cf. Section 5.2.2), while the noise standard deviation estimate will be provided to adjust the AutoPEEP detector.

AutoPEEP detector The AutoPEEP detector is the main core of the whole platform. Given a specified tolerance τ and the desired maximum false-alarm rate (level) γ, the AutoPEEP detector will decide whether an AutoPEEP is present or not for a given breath, on the basis of its observation $\mathbf{Y}_k$ and estimated parameters $\mathbf{p}_k, \hat{\sigma}$.

5.2.2 Detectors

Given tolerance τ and observation y_n of the noisy flow signal, the AutoPEEP detection is the testing of the null hypothesis $|f_{t_k}| \leqslant \tau$ against the alternative one $|f_{t_k}| > \tau$. The problem is then detecting signal deviation at critical instants t_k as presented in Section 4.1. The RDT framework introduced in Chapter 3 suggests optimal thresholding tests to such problem. Since the detection is carried out on end-expiration of consecutive breaths, both types of test shown in Section 4.1 can be used. As such, in this section, two AutoPEEP detectors are proposed. One — based directly on RDT — takes each of the breaths into account independently. The other — based on the sequential extension of RDT — is developed under the assumption that the state (AutoPEEP/NON-AutoPEEP) of the patient-ventilator interface is regular and remains the same within a certain number of breaths. With respect to the persistence of the phenomena and the breath rate, this assumption usually holds in practice.

Single-breath RDT-based AutoPEEP detector

The presence of an AutoPEEP is defined based solely on the final sample of the expiratory phase of each breath. However, since samples of flow signal in the expiratory phase exhibit a particular waveform, it is then expected that taking multiple samples into account will improve the detection performance. By introducing the so-called waveform vector, namely $\mathbf{p}_k$ as in Section 4.1.1, with dimension L, one can aggregate L samples at the end of the expiration to carry out a single decision for the breath under consideration. Let $\mathbf{Y}_k$ be the observation vector containing the last L samples of the expiratory phase of the k-th breath under consideration. $\mathbf{Y}_k$ is modeled as:

$$\mathbf{Y}_k = \mathbf{f}_k + \mathbf{X}_k$$

where $\mathbf{f}_k = \begin{bmatrix} f_{t_k-L+1} & \dots & f_{t_k-1} & f_{t_k} \end{bmatrix}^T$ is the flow signal vector and $\mathbf{X}_k \sim \mathcal{N}(0, \sigma^2 \mathbb{1}_L)$ is additive gaussian noise with standard deviation σ. Using the same factorization as in Section 4.1.1, vector $\mathbf{f}_k$ can be rewritten as:

$$\mathbf{f}_k = \mathbf{p}_k f_{t_k}$$

where $\mathbf{p}_k = \begin{bmatrix} p_1^{(k)} & p_2^{(k)} & \dots & p_L^{(k)} \end{bmatrix}^T$ is the waveform vector with $p_L^{(k)} = 1$. This vector $\mathbf{p}_k$ corresponds to the local form of the flow signal near the end of the expiratory phase. It is

also worth mentioning that this local waveform vector $\mathbf{p}_k$ is scale-invariant and depends mainly on the configuration of the interface, including the patient condition (pulmonary compliance, airway resistance, etc) and the ventilator settings (breath rate, relative timing of inspiration and expiration, etc). As long as the interface stays unchanged, the waveform vector remains almost the same regardless whether or not an AutoPEEP might occur. In practice, either $\mathbf{p}_k$ is known prior to the detection or it can be estimated from the observation using one of the methods proposed in Section 5.2.4.

To aggregate L observed samples into one decision for the considered breath, $\mathbf{Y}_k$ is projected onto the direction generated by $\mathbf{p}_k$ as suggested in Section 4.1.1 to obtain:

$$u_k = f_{t_k} + w_k \tag{5.1}$$

where $u_k = \mathbf{p}_k^T \mathbf{Y}_k / \|\mathbf{p}_k\|^2$, $w_k = \mathbf{p}_k^T \mathbf{X}_k / \|\mathbf{p}_k\|^2$ and $\|\mathbf{p}_k\|^2 = \mathbf{p}_k^T \mathbf{p}_k$ is the L_2-norm of waveform vector $\mathbf{p}_k$. By such processing, noise w_k follows normal distribution with zero mean and variance $\sigma_w^2 = \sigma^2 / \|\mathbf{p}_k\|^2$. According to [Billingsley, 1995, Theorem 27.4, p. 362] and equation (5.8), it can be proved that, even when the original noise is not gaussian, the resulting noise w_k tends to a normally distributed random variable, as long as L is large enough and the original noise samples are i.i.d (independent and identically distributed) (cf. Appendix C for the detailed proof). In practice, the i.i.d condition can be significantly relaxed. The aggregated observation u_k suffers from aggregated noise with lower standard deviation (i.e. $\sigma_w \leqslant \sigma$). Moreover, no information on the correlation among samples of noise vector $\mathbf{X}_k$ is required. The two hypotheses $h_0 : |f_{t_k}| \leqslant \tau$ and $h_1 : |f_{t_k}| > \tau$ are unchanged. The detection is thus carried out as follows:

$$\begin{cases} \text{If } |u_k| > \sigma_w \lambda_\gamma(\frac{\tau}{\sigma_w}) \text{ then we decide that there is AutoPEEP} \\ \text{Otherwise, the considered breath is labeled with Non-AutoPEEP.} \end{cases} \tag{5.2}$$

It should be noted that $\|\mathbf{p}_k\|$ increases with respect to the number L of samples. The noise standard deviation σ_w will thus decrease when more samples are taken into account. By reducing the noise standard deviation, the detection probability is improved while the false-alarm rate is always limited to the specified level γ. Theoretically, L is only limited by the time given for expiratory phase and the sampling rate of the data acquisition block. However, L must not be too long so that the local waveform vector can be considered stable and stays almost unchanged for a given number of breaths.

Sequential RDT-based AutoPEEP detector

This detector is a direct application of the so-called Sequential RDT presented in Section 4.1.2. Let us consider the flow signal observations of K consecutive breaths. For samples of each breath, the same projection technique as that used by the Single-breath

RDT-based AutoPEEP detector of the foregoing section yields the new observation: $u_k = f_{t_k} + w_k$ for $k = 1, 2, ..., K$ and $w_k \overset{iid}{\sim} \mathcal{N}(0, \sigma_w^2)$. By averaging over the considered K consecutive breaths — i.e. put $u_{1:K} = \frac{1}{K}\sum_{k=1}^{K} u_k$, $f_{t_{1:K}} = \frac{1}{K}\sum_{k=1}^{K} f_{t_k}$ and $w_{1:K} = \frac{1}{K}\sum_{k=1}^{K} w_k$ — we obtain the K-breath averaged-observation:

$$u_{1:K} = f_{t_{1:K}} + w_{1:K} \tag{5.3}$$

where averaged noise $w_{1:K}$ is gaussian with variance $\sigma_{w,K}^2 = \frac{\sigma_w^2}{K}$ strictly decreasing when more and more breaths are taken into account. Although the gaussianity of noise w_k — and, thus, that of noise $w_{1:K}$ — is asymptotically guaranteed by the projection technique when the number L of samples is sufficiently large, it should be noticed that the averaging over K breaths could guarantee the gaussianity of K-breath averaged-noise $w_{1:K}$ without requiring the gaussianity of aggregated-noise w_k, as long as w_k is i.i.d and K is itself large enough. Since K increases in sequential framework, it can be said that the gaussianity of noise can be assumed with more and more confidence when observations are being collected.

Assuming that the patient-ventilator interface — and, therefore, the true label (AutoPEEP/NON-AutoPEEP) — remains the same for K consecutive breaths under consideration, the AutoPEEP detection for these breaths amounts to determining whether or not the average end-expiration flow $f_{t_{1:K}}$ exceeds the specified tolerance τ. Given some desired level γ, the from-above test for this problem is:

$$\mathcal{T}_{\lambda_{1:K}^{(h)}}(u_{1:K}) = \begin{cases} 1 & \text{if} \quad |u_{1:K}| > \lambda_{1:K}^{(h)} \\ 0 & \text{if} \quad |u_{1:K}| \leqslant \lambda_{1:K}^{(h)} \end{cases}$$

for testing $[h_0 : |f_{t_{1:K}}| \leqslant \tau]$ against $[h_1 : |f_{t_{1:K}}| > \tau]$ and the from-below test is:

$$\mathcal{T}'_{\lambda_{1:K}^{(\ell)}}(u_{1:K}) = \begin{cases} 1 & \text{if} \quad |u_{1:K}| \leqslant \lambda_{1:K}^{(\ell)} \\ 0 & \text{if} \quad |u_{1:K}| > \lambda_{1:K}^{(\ell)} \end{cases}$$

for testing $[h'_0 : |f_{t_{1:K}}| > \tau]$ against $[h'_1 : |f_{t_{1:K}}| \leqslant \tau]$. The two associated thresholds are computed as:

$$\lambda_{1:K}^{(h)} = \sigma_{w,K}\lambda_\gamma(\tau/\sigma_{w,K})$$

and

$$\lambda_{1:K}^{(\ell)} = \sigma_{w,K}\lambda_{1-\gamma}(\tau/\sigma_{w,K})$$

where $\lambda_{1:K}^{(h)} > \lambda_{1:K}^{(\ell)}$ for any $0 < \gamma < 0.5$. The proposed dual-threshold test is then $\mathcal{T}''_{[\lambda_{1:K}^{(\ell)}, \lambda_{1:K}^{(h)}]}(u_{1:K})$ for testing $[h_0 : |f_{t_{1:K}}| \leqslant \tau]$ against $[h_1 : |f_{t_{1:K}}| > \tau]$. The decision suggested by this dual-threshold test on the basis of these K breaths is given by:

$$[d(u_{1:K})] = \begin{cases} \textbf{if} \quad |u_{1:K}| > \lambda_{1:K}^{(h)} \quad \textbf{then} & \text{AutoPEEP is present } (h_1) \\ \textbf{if} \quad |u_{1:K}| \leqslant \lambda_{1:K}^{(\ell)} \quad \textbf{then} & \text{NON-AutoPEEP } (h_0) \\ \textbf{otherwise} & \text{Not decided yet. The decision} \\ & \text{is delayed until next observation.} \end{cases} \tag{5.4}$$

Summarizing, as long as the breaths take place one after another, the observations are sequentially acquired and the AutoPEEP detection is then carried out sequentially as well. The process is the same as that described in Section 4.1.2. More specifically, it begins with the detection on the first observation u_1 using $[d(u_{1:1})]$. Depending on the result returned by $[d(u_{1:1})]$, either the process will stop with a label for the observed breath or the decision must be delayed until the arrivals of the next observations $u_2, u_3, \ldots$. The process will be completed and restarted as soon as either a decision is made — i.e. there exists a value K so that $|u_{1:K}| \notin (\lambda^{(\ell)}_{1:K}, \lambda^{(h)}_{1:K}]$ — or the maximum number M of breaths have been observed and the hard decision:

$$[d_h(u_{1:M})] = \begin{cases} \textbf{if} \quad |u_{1:M}| > \lambda^{(h)}_{1:K} \quad \textbf{then} \quad \text{AutoPEEP is present } (h_1) \\ \textbf{if} \quad |u_{1:M}| \leqslant \lambda^{(h)}_{1:K} \quad \textbf{then} \quad \text{NON-AutoPEEP } (h_0) \end{cases} \tag{5.5}$$

must be used. This number M of breaths is specified by the clinician regarding the breath rate and the stability of the patient-ventilator interface under consideration. It must be selected so that the state (AutoPEEP/NON-AutoPEEP) of the interface still remains constant during these M observations and the maximum delay-to-decision is respected. In our experimental settings, we use $M = 10$, which corresponds to about 30 seconds of signal recording in the usual case with a breathing frequency of 20 [breaths/min].

5.2.3 Phase change detection

Since the detection of AutoPEEP is performed on the basis of the flow samples at the end of the expiratory phase of each breath, it is required that the instant where expiratory phase ends can be precisely retrieved. The Phase change detection/segmentation block in Figure 5.2 is included to accomplish such a task. More specifically, its main role is to provide the exact instant of end-expiration for each considered breath. This can be achieved by detecting the change in flow signal y_n from the expiratory phase of the current breath (negative values) to the inspiratory phase of the next breath (positive values) (c.f. Figure 5.1). In practice, the signal perturbation caused by unavoidable noise or any other unexpected physical effects might get involved and may bias the detection. In such cases, a smoothed version of the signal can be used instead of the observed signal itself. To be simple, the *moving average smoothing* (SMA) method can be considered:

$$\bar{y}_n = \text{SMA}(y_n) = \frac{1}{2h+1} \sum_{i=n-h}^{i=n+h} y_i$$

where $2h + 1$ is the length of the moving window.

It is worth mentioning that each single breath can be divided into three different respiratory phases — namely Inspiration, Pause (when it is specified by clinician) and

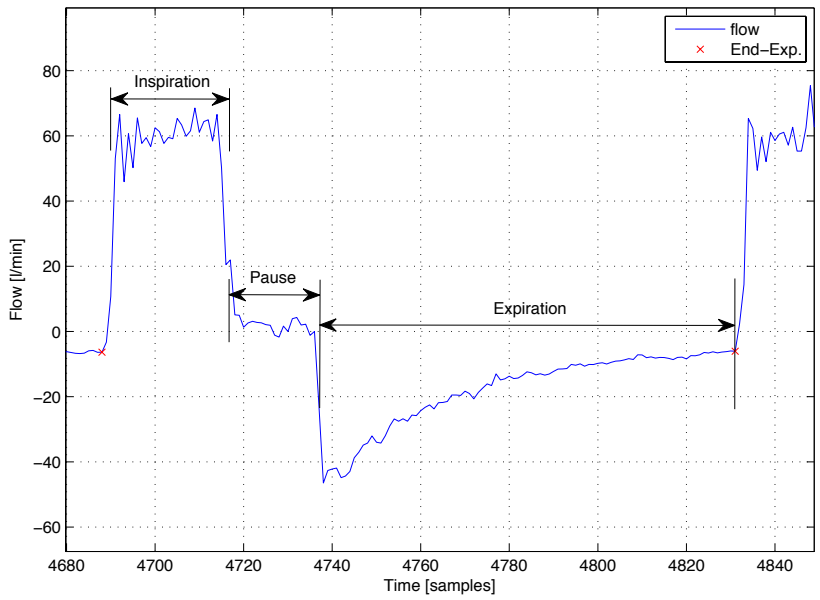


Figure 5.3 — Flow signal during different phase of a single breath. Three consecutive respiratory flow are usually observed: Inspiration, Pause and Expiration.

Expiration — as shown in Figure 5.3. During each of these phases, the flow signal exhibits a particular waveform. These waveforms depend on both the patient parameters and the mechanical ventilatory specifications given by clinician. In particular, during the inspiratory phase, the flow signal waveform depends mainly on the ventilator parameters specified by the clinician. It can be constant (in Constant Flow mode), first-order polynomial (in either Accelerating Flow or Decelerating Flow mode) or has a sinusoidal form (in Sinusoidal Flow mode). In most cases, the Constant Flow is practiced; therefore, the flow signal remains nearly constant with a positive value which represents the inhaling direction of the air-flow. During a respiratory pause (if it is present), since both inhaling and exhaling air valves are closed, the observed flow signal is nearly constant at null value. On the contrary, during expiratory phase, the functioning of the whole system depends mainly on passive action of the patient undergoing ventilatory support. The expiratory flow signal waveform is thus determined by internal parameters of the patient's respiratory system, including airway, lung, chest wall, etc. Due to the resistance of the airway and the elasticity of patient lungs/chest walls, the flow signal during this expiratory phase exhibits an exponential waveform of the same type as an RC circuit (see Section 5.2.4 for more details).

Given flow signal y_n or its smoothed version $\bar{y}_n$, the observed signal is regular during each respiratory phase except for transitions from one phase to another where irregularities are presents. The detection of end-expirations, which is a phase transition from expiration to inspiration, is thus a direct application of Problem 2 **[Chg.]** mentioned above (cf. Section 4.2). The approach based on a transformation (for example, the wavelet transform), the universal threshold and the RDT as presented in Section 4.2

can thus be used to address such a detection application. For the flow signal in Figure 5.1, the end-expiration detection result is given as in Figure 5.4. In this example,

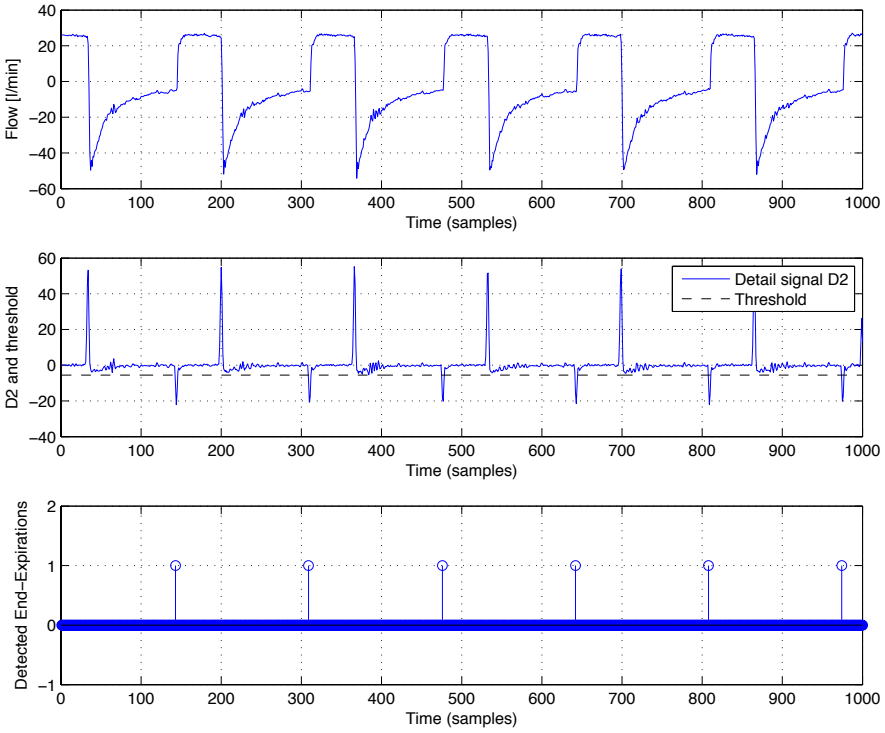


Figure 5.4 — End-Expiration Detection using Wavelet transform. This figure illustrates the detection of end-expirations based on respiratory flow signal: (top) respiratory flow curve obtained from a patient, (middle) signal in the level-2 detail band of the wavelet transform coefficients and the calculated detection threshold, (bottom) detection result, where 1's (peaks) represent end-expirations.

the stationary Haar wavelet transform was used and the level-2 detail band of wavelet coefficients was taken into account. It should also be noticed that, since a peak is only one point, the results of the thresholding test should be post-processed in such a way that consecutive 1's are removed. In particular, in case of consecutive decisions equal to 1, only the first one will be kept. Apparently, the one corresponding to the highest absolute coefficient could also be chosen instead of the first one. However, keeping the first one makes it possible to reduce the detection delay and the computational complexity. Last but not least, since the transition from expiratory to inspiratory phase is a change from negative to positive flow, end-expirations are negative peaks.

It could be noted that the noise standard deviation in detail band wavelet coefficient relates to that of the original observation. Therefore, it can be inferred from the noise standard deviation estimator given in Section 5.2.4 below. Any other estimation methods can be used.

Although there might exist many other approaches to accomplish such end-expirations detection in literature, in this section, we consider the approach based

on transformation, universal threshold and the RDT for its simplicity and robustness. In practice, any other end-expiration detection which might yield better performance can be applied.

It could be reminded that the main purpose of this functional block is to provide end instants of expiratory phases for the detection of AutoPEEP. However, in addition to end-expiration detection, the segmentation of each breath into different phases is profitable in various tasks, including noise estimation and waveform vector regression. It also provides much information for the detection of various types of patient-ventilator asynchrony related to triggering instants and respiratory phase timing. In this regard, the approach proposed in this work can be used to provide such segmentation. Furthermore, based on fact that flow signal exhibits specific waveforms in different respiratory phases, other approaches such as Hidden Markov Models, Segmental Models could also be investigated.

5.2.4 Estimations

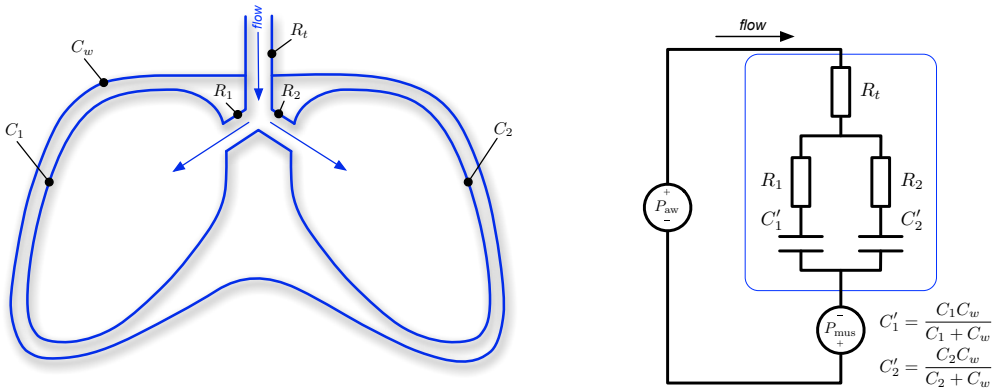
As shown in Section 5.2.2, some parameters, including the waveform vector ($\mathbf{p}_k$) and the unknown noise standard deviation (σ) must be given to the AutoPEEP detector prior to performing a decision. These parameters are supposed to be known in practice. Otherwise, they can be estimated via observation of the flow signal under consideration. Since only the portion of the signal corresponding to expiratory phase is of interest, the segmentation given by the Phase change detector might also be used. In what follows, two estimations — one is for the so-called waveform vector $\mathbf{p}_k$ and the other is for the standard deviation of unknown noise — will be addressed.

Waveform regression to compute $\mathbf{p}_k$

With regard to Section 5.2.2, the waveform vector $\mathbf{p}_k$ is the key that makes it possible to aggregate multiple end-expiration flow samples into one decision, and thus improves the probability of detection by reducing the noise effect on the decision. In practice, thanks to the particular waveform of the flow signal during expiratory phase of a breath, this vector $\mathbf{p}_k$ can be computed via a regression task.

Indeed, during the expiratory phase of a breath, the inhaling valve is closed to stop the ventilator from pumping air into the patient lung and the exhaling valve is left open for the patient to exhale on his/her own. During this respiratory phase, the ventilator works based exclusively on the passive response of the patient undergoing mechanical ventilation; and, therefore, the waveform exhibited by the flow signal depends mainly on the mechanical characteristics of his/her respiratory system. Basically, human respiratory system consists of the pulmonary system — which includes the lungs and the

airways — and the chest wall — which comprises breath producing mechanical structures such as the rib cage and the respiratory muscles. These organisms present passive mechanical characteristics related to breathing, including resistances and compliances, as shown in Figure 5.5(a). As a result, the mechanical response of the patient's respiratory system can be modeled by an electrical circuit with resistors and capacitors in Figure 5.5(b). In these figures, the model with two compartments was considered.



(a) Human respiratory system parameters (b) Equivalent circuit

Figure 5.5 — Human respiratory system parameters and the equivalent electrical circuit with Rs, Cs

With respect to the equivalent electrical circuit, the flow signal during the expiratory phase of a breath can be modeled by an exponential form as follows:

$$y(t) = C_0 - \phi e^{-\mu t}, \tag{5.6}$$

with $\phi > 0$ and $\mu > 0$, regardless the absence/presence of AutoPEEP. This model can be used to estimate the reference waveform at the ending portion of the expiration using a nonlinear robust regression method. Given a set of N data points $\{(t_i, y(t_i)), i = 1..N\}$ where $y(t_i)$ is the observation at instant t_i, the non-linear robust regression aims at solving the least square problem:

$$(\hat{C}_0, \hat{\phi}, \hat{\mu}) = \underset{C_0,\phi,\mu}{\arg\min} \sum_{i=1}^{N} \xi_i \left[y(t_i) - (C_0 - \phi e^{-\mu t_i}) \right]^2 \tag{5.7}$$

where ξ_i is weight given to the observation at t_i. The introduction of weight vector $[\xi_1, \xi_2, .., \xi_N]$ makes it possible to reduce the influence of outliers onto the final result. Given optimal set of parameters $(\hat{C}_0, \hat{\phi}, \hat{\mu})$, the regressed values are then calculated as $\hat{y}(t_i) = (\hat{C}_0 - \hat{\phi} e^{-\hat{\mu} t_i})$. For implementation, one could consider employing the Matlab *nlinfit* routine to carry out such a regression task. This routine uses a weighted version of the Levenberg-Marquardt algorithm [Seber and Wild, 2003] to solve the non-linear least squares problem (5.7). Coefficients ξ_i given to observations y_i are iteratively updated with respect to corresponding residues $|y(t_i) - (C_0 - \phi e^{-\mu t_i})|$ $(i = 1..n)$

to downweight the outliers and therefore reduce their effects on the final regression curve. Figure 5.6 shows an example of the flow signal at the end of the expiratory phase and the regression resulting from the aforementioned non-linear robust method. The signal has been shown to be well-fitted by the model function (5.6).

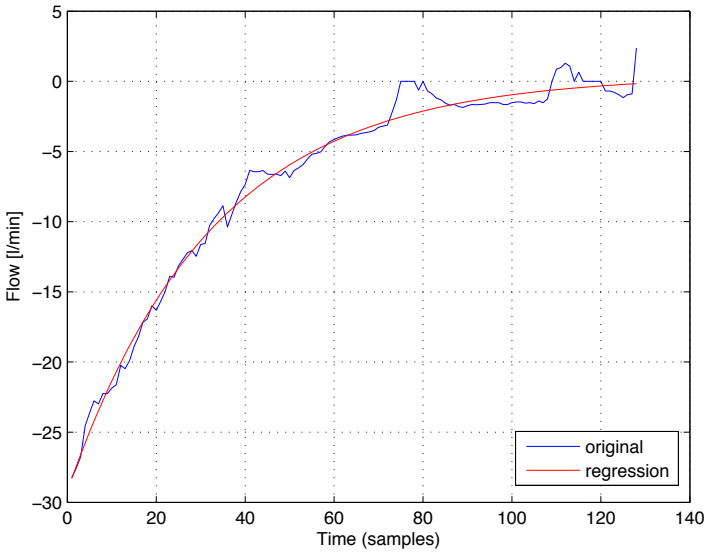


Figure 5.6 — Fitness of the model function. An example of the flow signal at the end of an expiratory phase with its regression curve using the model function in (5.6). The result firmly shows the relevance of the considered model function to the regression task.

It must be noticed that, although only L samples are enough for calculating the L-dimensional waveform vector $\mathbf{p}_k$, more samples should be used to achieve a better regression curve. Let L_{ext} ($L_{\text{ext}} \geqslant L$) be the number of observation samples to be used. L_{ext} is only limited by the length, namely L_E (in samples), of the observed flow signal in expiratory phase, i.e. $L_{\text{ext}} \leqslant L_E$. Regarding the transition between different respiratory phases, samples at the beginning of expiration are very sensitive to transition and may bias the regression. Therefore, only a proportion of the L_E samples of the expiratory phase should be taken into account:

$$L_{\text{ext}} = \lfloor \alpha L_E \rfloor$$

where $\lfloor . \rfloor$ is the *floor* function and α ($0 < \alpha < 1$) specifies the percentage of expiratory samples to be used. Proportion α must be chosen so that $L_{\text{ext}} = \lfloor \alpha L_E \rfloor \geqslant L$. Furthermore, to avoid the border effect, one might consider an additional weighting scheme that puts more weight on the middle samples than on the side ones. Figure 5.7 shows the regression on end-expiration samples of a flow signal recorded from a patient undergoing mechanical ventilatory support. In this example, α is set to 0.75 to avoid the transition effects at the beginning of the expiratory phase. Since it is not so crucial in the situation experienced in this application, all L_{ext} samples in the observing win-

dow are considered with the same importance. No additional weighting strategy was employed.

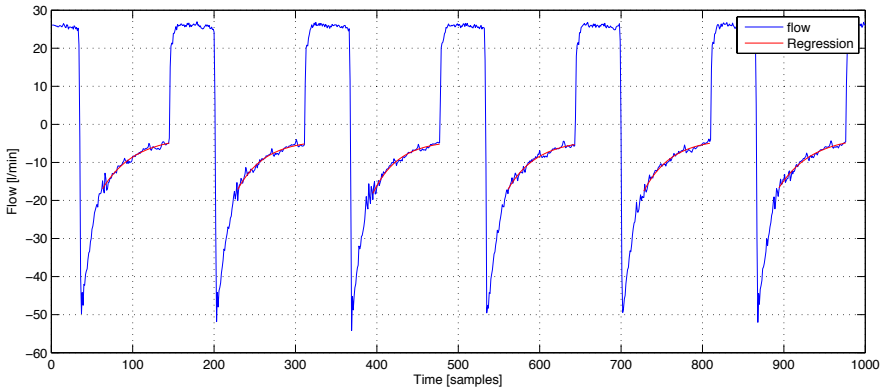


Figure 5.7 — An example of the flow signal on real patient with end-expiration regression results.

Given the regression values $[\hat{y}_{t_k-L_{\text{ext}}+1}, \hat{y}_{t_k-L_{\text{ext}}+2}, ..., \hat{y}_{t_k}]$ at the ending portion of the flow signal in expiratory phase, the last L values are used to compute the estimate $\hat{\mathbf{p}}_k$ for the considered breath as follows:

$$\hat{\mathbf{p}}_k = \frac{1}{\hat{y}_{t_k}} \begin{bmatrix} \hat{y}_{t_k-L+1} \\ \hat{y}_{t_k-L+2} \\ ... \\ \hat{y}_{t_k} \end{bmatrix} \tag{5.8}$$

This vector is scale invariant and depends mainly on the passive mechanical properties of the patient's respiratory system, more precisely, the resistance of the airways and the compliances presented by the lungs and the chest walls.

According to Section 5.2.2, for aggregating multi-samples into a single decision, waveform vector $\mathbf{p}_k$ concerns solely the current (k-th) breath. Vector $\mathbf{p}_k$ can then be estimated, as shown above, based exclusively on observed flow signal samples of this breath. However, in practice, the resistance and the compliance introduced by the patient's respiratory system do not vary much during the monitoring. Therefore, the waveform vector remains almost the same from one breath to another. This fact make it sensible to use estimates from previous breaths so as to improve the estimation of $\mathbf{p}_k$ for the current k-th one. In this respect, the following strategies can be considered to compute the waveform vector estimate to be used in AutoPEEP detectors:

Static waveform vector:
In case the interface — in particular the passive mechanical properties from the patient side — is stable, the waveform vector can be computed on the basic of the first N_{ref} breaths with the validation of the clinician who regularly perform a verification/tuning session after a specific time of automatic monitoring. These

breaths are considered as the reference for the rest of the observed signal. Given N_{ref} waveform vectors estimated from the referenced breaths, the vector to be used is that given by:

$$\bar{\mathbf{p}} = \frac{1}{N_{\text{ref}}} \sum_{k=1}^{N_{\text{ref}}} \hat{\mathbf{p}}_k$$

This waveform vector will be constantly employed until it is updated due to some modifications of the ventilator parameters given by the clinician during the next tuning/verification session. It is also thinkable to carry out the update on a regular time basis.

It is worth mentioning that, in practice, when other types of asynchrony are taken into account, it is then expected that one or more references are available for the monitoring task. The static waveform vector is suitable in such cases. Moreover, it requires no more regression/estimation during the monitoring, therefore, reduces real-time computation task.

Dynamic waveform vector:
The waveform vector to be used is the one estimated from the current breath:

$$\bar{\mathbf{p}}_k = \hat{\mathbf{p}}_k$$

No correlation with previous breaths is exploited. Contrary to the static waveform vector, this strategy is suitable for situations in which the interface varies more than expected and the variation of waveform vector estimate from one breath to another is beyond the accepted range. The regression/estimation is, however, required for every new breath acquired.

Adaptive waveform vector:
In this strategy, the waveform vector is updated every time a new breath is observed. Previous estimates are taken into account with a forgetting factor μ such that $0 < \mu < 1$:

$$\bar{\mathbf{p}}_k = \frac{1-\mu}{1-\mu^k} \sum_{i=1}^{k} \mu^{k-i} \hat{\mathbf{p}}_i$$

This forgetting factor controls how estimates from previous breaths involve the waveform vector used for the current (k-th) one. The choice of μ should reflect the variation among waveform vector estimates and, thus, the stability of the interface. The adaptive waveform vector tends to the dynamic one when $\mu \to 0$ and amounts to the static solution with $N_{\text{ref}} = k$ when $\mu \to 1$.

Estimation of the noise standard deviation

Noise is usually unknown in practice. As long as the noise standard deviation is concerned, it can be estimated from the observation. In this study, two solutions are

considered: one based directly on the result obtained by waveform regression, whereas the other is based on an estimation from the wavelet coefficients of the observed flow signal.

Estimation via regression residues:
By using the function in (5.6) as the model, the regressed curve $\hat{y}(t)$ provides an estimation for the flow signal in expiratory phase of the considered breath. The regression residues can be considered as noise. Therefore, they can be used to directly estimate the required noise standard deviation. In particular, we have:

$$\hat{\sigma}_k = \frac{1}{L_{\text{ext}} - 1}\sqrt{\sum_{i=t_k-L_{\text{ext}}+1}^{t_k}(y_i - \hat{y}_i)^2}$$

for the k-th breath. To aggregate $\hat{\sigma}$ from $\hat{\sigma}_k$, the similar strategies as those proposed for the waveform vector can be considered. The choice of aggregating strategies depends on the variation of noise with respect to the duration of one breath.

It should be noticed that, strictly speaking, the regression includes both original noise and regression error. The residue standard deviation estimate is than greater than or equal to that of actual noise. As the result, the constraint on the probability of false-alarm will alway be respected.

Estimation via wavelet coefficients:
This approach investigates the sparsity of the noisy signal in the transform domain. Indeed, studies on nonparametric estimation based on Wavelet Shrinkage have shown that most of the wavelet coefficients obtained from the first level wavelet decomposition of a piecewise smooth signal are of very small amplitude. Only a small number of these wavelet coefficients, which correspond to signal, are of higher amplitude [Donoho and Johnstone, 1994]. This fact allows the use of robust estimators on the wavelet coefficients to provide noise estimation.

For reference, one could consider the MAD (median absolute deviation) [Hampel, 1974, Rousseeuw and Croux, 1993] to accomplish such a task. The method is usual and can be found in [Hampel, 1974, Rousseeuw and Croux, 1993, Donoho and Johnstone, 1994]. We recall it here for readiness sake. Let $c_1, c_2, ...c_N$ be the wavelet coefficients obtained from the first level discrete wavelet decomposition of an N-sample segment of the flow signal y. The estimate $\hat{\sigma}_{\text{MAD}}$ of σ is then provided by:

$$\hat{\sigma}_{\text{MAD}} = b \times \text{med}_i|c_i - \text{med}_j c_j|$$

where $b \approx 1.4826$. As long as the noise is central, white and gaussian, the formula is simplified to:

$$\hat{\sigma}_{\text{MAD}} = b \times \text{med}_i|c_i|$$

knowing that $\text{med}_i c_i = 0$.

In [Pastor and Socheleau, 2012], another robust estimator was proposed, namely the *d-dimensional adaptive trimming estimator* (DATE). The method is summarized as follows. Let $c_{(1)}, c_{(2)}, ..., c_{(N)}$ be sequence of wavelet coefficients $c_1, c_2, ...c_N$ sorted by increasing magnitude. Put $m_{\min} = \frac{N}{2} - \sqrt{\frac{N}{4(1-Q)}}$ where $Q = 0.95$. Let m be the smallest integer, $m_{\min} \leqslant m \leqslant N$, such that:

$$|c_{(m)}| \leqslant 2.7238 \times \frac{1}{m} \sum_{k=1}^{m} |c_{(k)}| < |c_{(m+1)}|$$

If such an integer m does not exist, set $m = m_{\min}$. The estimate $\hat{\sigma}_{\text{DATE}}$ of σ is then provided by:

$$\hat{\sigma}_{\text{DATE}} = 1.2533 \times \frac{1}{m} \sum_{k=1}^{m} |c_{(k)}|$$

It has been shown in [Pastor and Socheleau, 2012] that this estimator outperforms the MAD when the number of outliers increases. The DATE can be employed as an alternative to the MAD mentioned above when the proportion of outliers is between 0.3 and 0.5 (cf. [Pastor and Socheleau, 2012]). For the cases considered in this work, because the number of large wavelet coefficients pertaining to signal remains small, the two estimators yield similar performance. The MAD estimator is thus adopted for its lower complexity and higher rapidity.

In term of complexity, the estimation via wavelet coefficient requires more computational resource due to the transformation of observed signal into the wavelet domain. Moreover, it requires that the signal is processed by packets. The size N of each packet must be neither too short for the estimation to be consistent nor too long so as the detection can be carried out in real time. On the contrary, the estimation from regression is direct and it counts only observation samples in expiratory phase which is of interest. This is an advantage since it could be happened in practice that the noise is different in each respiratory phase.

5.3 Detection performance assessment

As mentioned in Section 5.1, the detection performance of the proposed framework was assessed in three different settings. The proposed AutoPEEP detectors were first evaluated via simulations on computer. Synthesized flow data were generated for this purpose. Secondly, a respiratory system analog was pneumatically connected to a currently used mechanical ventilator to carry out ventilatory support emulations. The proposed platform was then tested with the emulated data derived from the respiratory system analog in use. For further assessment, real data recorded during the practice

of mechanical ventilator support on patients were retrospectively analyzed. In what follows, the three settings will be presented. The results will also be reports. These results would give an idea on the performance of the proposed framework before a further prospective study in a real-time monitoring configuration can be organized.

5.3.1 Simulations

In this assessment setting, simulations were implemented on computer to illustrate the detection performance of the proposed detection algorithms, i.e. the Single Breath RDT-based AutoPEEP detector and the Sequential RDT-based one (cf. Section 5.2.2). The flow signal was first synthesized (see Figure 5.8 for an example of such flow signal). For each breath, L end-expiration flow signal samples were generated. The waveform

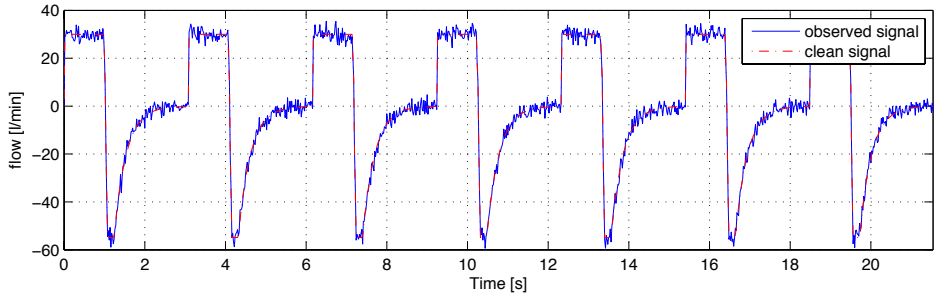


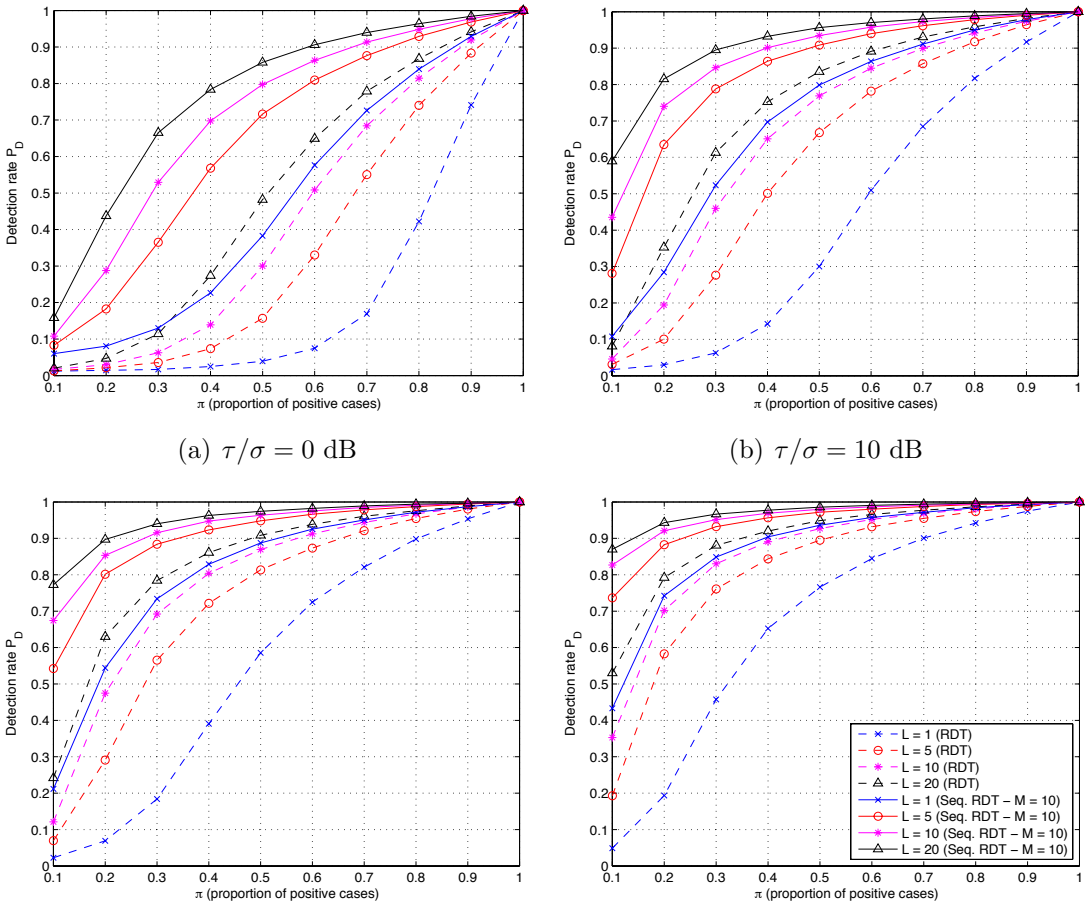
Figure 5.8 — An example of synthesized flow signal. In this example, noise standard deviation was set $\sigma = 2$ [l/min] which yields $\tau/\sigma = 0$dB

vector was supposed to be known and set to $\mathbf{p}_k = \mathbf{p} = [1, 1, ..., 1]^T$. It is worth mentioning that, by construction, $|p_1| \geqslant |p_2| \geqslant ... \geqslant |p_L| = 1$ and, as a result, $\sigma_w = \frac{\sigma}{\|\mathbf{p}_k\|} \leqslant \frac{\sigma}{\sqrt{L}}$. The equality happens when and only when $p_i = 1$ for all $i = 1..L$. With regard to standard deviation σ_w of the aggregated noise, by setting $\mathbf{p}_k = \mathbf{p} = [1, 1, ..., 1]^T$, we actually considered the worst case where $\|\mathbf{p}_k\|^2 = L$ and $\sigma_w = \frac{\sigma}{\sqrt{L}}$. In practice, when $\mathbf{p}_k \neq \mathbf{p} = [1, 1, ..., 1]^T$, the detection performance must be better than that given in this setting. The values of f_{t_k} were randomly and uniformly generated between 0 and $-\frac{\tau}{1-\pi}$, where π is the proportion of positive cases (AutoPEEP). The tolerance was empirically set to $\tau = 2$ [l/min] thanks to experts' experience of the domain and a level $\gamma = 0.01$ was specified by clinician. The same simulations with a level $\gamma = 0.05$ were also carried out.

For the sequential RDT-based detector, M was set to 10 [breaths]. With a usual breath rate of 20 [breaths/min], this value of M corresponds to about 30 seconds of signal observation. This is also the maximum delay that a decision can be postponed until enough evidence can be accumulated.

To visualize the detection performance, the ROC (Receiver Operation Curve) is usually plotted to illustrate how the detection rate $\mathrm{P_d}$ varies with respect to the false-

alarm rate P_{fa}. However, in detection framework similar to that presented by Neyman-Pearson, the false-alarm rate P_{fa} is always restricted to the specified value γ. Therefore, it is more meaningful to plot the detection rate P_d versus different values of π, namely the *detection curve*, than to present the usual ROC.



(a) $\tau/\sigma = 0$ dB

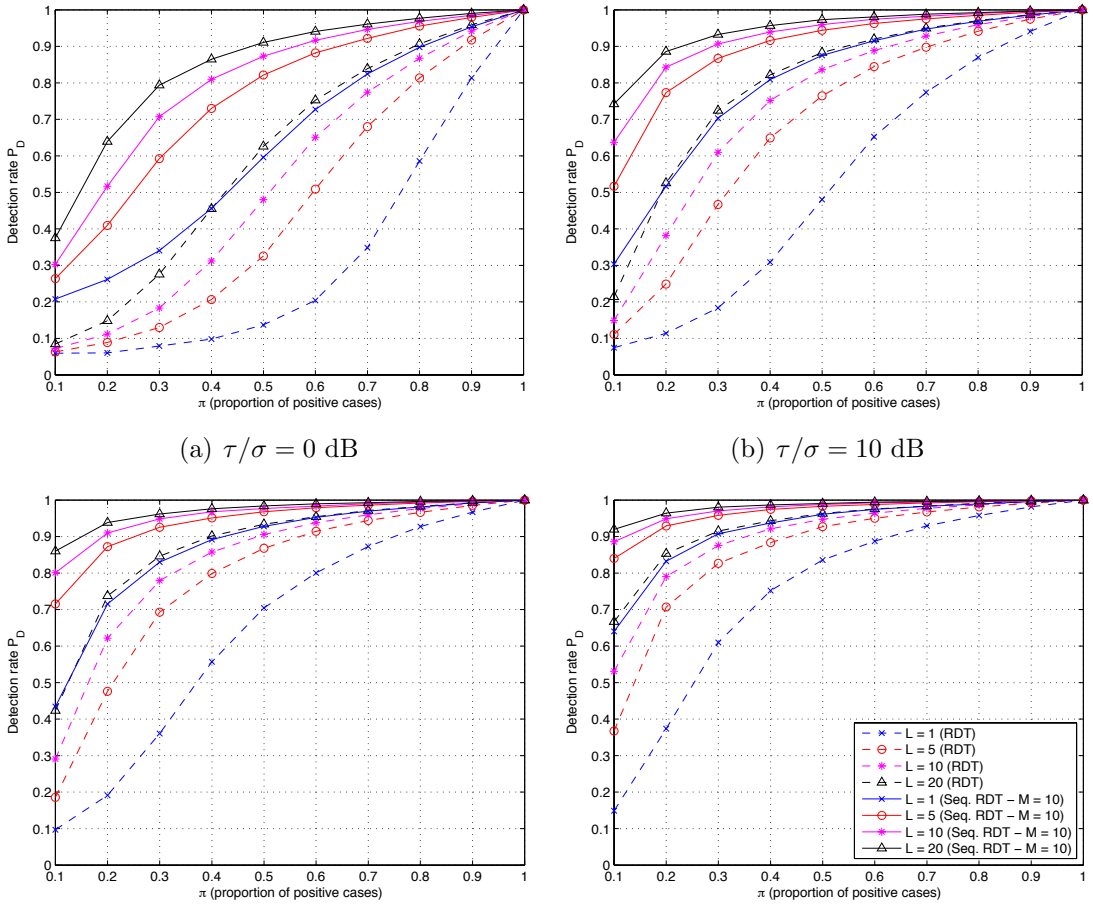
(b) $\tau/\sigma = 10$ dB

(c) $\tau/\sigma = 15$ dB

(d) $\tau/\sigma = 20$ dB

Figure 5.9 — Detection curves yielded by the two proposed AutoPEEP detectors with different noise levels. The simulations were carried out with $N = 10000$ breaths, tolerance $\tau = 2$ [l/min] and level $\gamma = 0.01$. With the extension of RDT in a sequential framework, the resulting detector yields a significant improvement in detection rate while the false alarm is still limited to the specific value γ.

In Figure 5.9 and Figure 5.10, detection curves are given with different noise levels and different values of L. It could be mentioned that, the detection rate is significantly improved when more samples are aggregated. In practice, the choice of L depends on the expiratory time given and the sampling rate of the flow sensor to be used. For example, a G5 ventilator (Hamilton Medical, Bonaduz, Switzerland) provides flow data stream at a rate of 50 [samples/s]. Therefore, in a real-world setting, when it is configured with a usual breath rate of 20 [breaths/min] and the inspiratory-to-expiratory time ratio $(I : E)$ equals $1 : 2$, the value of L samples to be aggregated can be set as high



(a) $\tau/\sigma = 0$ dB

(b) $\tau/\sigma = 10$ dB

(c) $\tau/\sigma = 15$ dB

(d) $\tau/\sigma = 20$ dB

Figure 5.10 — Detection curves with level $\gamma = 0.05$. The simulations were also carried out with $N = 10000$ breaths and tolerance $\tau = 2$ [l/min]. With a more relaxed constraint on the false-alarm probability, the higher detection rate is obtained.

as $L = 100$, which will yield a very good detection performance. Of course, the lower the noise level, the better the detection. In this respect, the Sequential RDT-based detector also showed higher detection rate while still keeping the false-alarm rate below the specified value γ. It could also be noticed that, as long as the constraint on the false-alarm probability is relaxed by increasing the level γ, the higher detection rate is obtained.

5.3.2 Emulations with a respiratory system analog

The proposed AutoPEEP detection framework was then tested in a more realistic setting in which the interface between a ventilator and a computerized respiratory system analog was established. The emulation was arranged as shown in Figure 5.11. In these experiments, the testbed was constituted by a G5 ventilator (Hamilton Medical,

Bonaduz, Switzerland) pneumatically linked to the ASL5000 computerized respiratory system analog (Ingmar Medical Ltd., Pittsburgh, PA, USA), making it possible to modify respiratory mechanics. The interface is totally configurable on both the ventilator side and the artificial respiratory patient end. A computer is connected to the ASL5000 via the RS232 communication protocol to, in the first step, set up its mechanical parameters and to, in the later step, record the respiratory curves output from its internal sensors for further processing. The proposed test bed allows us to mimic various scenarios with patients varying from neonatal to adult and in any health condition.

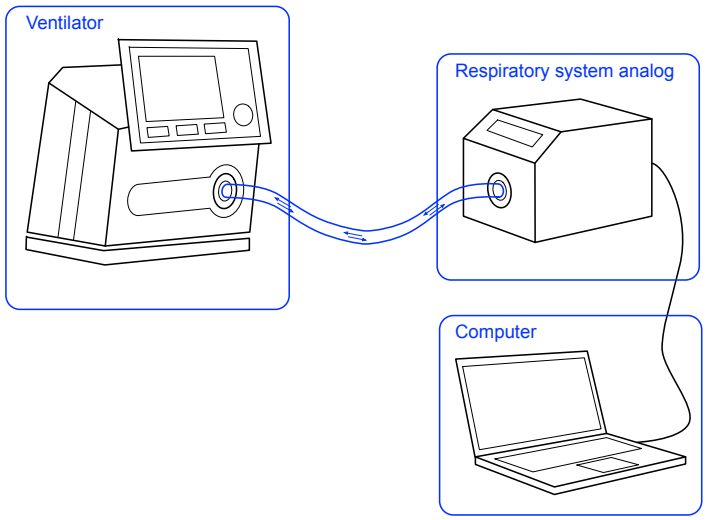


Figure 5.11 — Emulation testbed. In this setting, a computerized respiratory system analog, which can mimic any patient, is pneumatically connected to a currently used ventilator to undergo mechanical ventilatory support. A computer is linked to the artificial respiratory patient to parameterize its mechanical characteristics and to obtain the recorded flow signal.

For AutoPEEP detection performance assessment, the mechanical ventilatory support is carried out with thirteen different sets of parameters (cf. Table 5.1) for both the respiratory system emulator and the ventilator, which correspond to various practical situations. The empirical tolerance $\tau = 2$ [l/min] issued by experts was employed again. With respect to this tolerance, among the 13 settings, 7 cases were reported as AutoPEEP and the other 6 cases were labeled as NON-AutoPEEP, thanks to an independent clinical analysis from the Intensive Care unit of Brest University Hospital, Brest, France. The detection was performed on the basis of the flow signal captured by the sensor integrated in the ASL5000 respiratory system analog. For each case, about 1.5 minute of the signal flow was recorded. The corresponding number of breaths varied from 13 to 34, depending on the parameters that were used on the ventilator side. In total, 323 breaths were recorded. For both the proposed AutoPEEP detectors, the dynamic waveform vector was employed for its simplicity and flexibility. Level γ was set to 0.01. The detection results are detailed in Table 5.1.

Id	Parameters		True	N. of	Det. by RDT [c]			Det. by Seq. RDT [c]		
	Ventilator [a]	Lung model [b]	Label	breaths	Pos	Neg	Label	Pos	Neg	Label
1	PEP=0, Vt=500, Fbr=15, P=0, I:E=1:2	C=80, R=5	Neg	21	0	21	Neg	0	21	Neg
2	PEP=0, Vt=500, Fbr=15, P=0, I:E=1:2	C=30, R=5	Neg	20	0	20	Neg	0	20	Neg
3	PEP=0, Vt=500, Fbr=25, P=0, I:E=1:2	C=80, R=5	Pos	33	33	0	Pos	33	0	Pos
4	PEP=0, Vt=500, Fbr=25, P=0, I:E=1:1	C=80, R=5	Pos	34	34	0	Pos	34	0	Pos
5	PEP=0, Vt=300, Fbr=20, P=0, I:E=1:2	C=80, R=5	Neg	27	0	27	Neg	0	27	Neg
6	PEP=0, Vt=500, Fbr=12, P=0, I:E=1:2	C=80, R=5	Neg	16	0	16	Neg	0	16	Neg
7	PEP=0, Vt=500, Fbr=20, P=15, I:E=1:3	C=80, R=5	Neg	27	0	27	Neg	0	27	Neg
8	PEP=5, Vt=500, Fbr=20, P=0, I:E=1:3	C=80, R=5	Neg	27	0	27	Neg	0	27	Neg
9	PEP=5, Vt=500, Fbr=20, P=0, I:E=1:2	C=120, R=10	Pos	27	27	0	Pos	27	0	Pos
10	PEP=0, Vt=700, Fbr=20, P=0, I:E=1:2	C=120, R=10	Pos	27	27	0	Pos	27	0	Pos
11	PEP=0, Vt=700, Fbr=20, P=0, I:E=1:6	C=120, R=10	Pos	24	24	0	Pos	24	0	Pos
12	PEP=0, Vt=700, Fbr=20, P=0, I:E=1:1	C=120, R=10	Pos	27	27	0	Pos	27	0	Pos
13	PEP=0, Vt=700, Fbr=20, P=0, I:E=1:2	C=140, R=25	Pos	13	13	0	Pos	13	0	Pos

[a] Ventilator parameters include: Positive Expiratory Pressure PEP [cmH_2O], air volume Vt [ml], frequency Fbr [breaths/min], pause time P [%], Inspiratory to expiratory time ratio I:E.

[b] Lung model parameters include: compliance C [ml/cmH_2O] and resistance R [cmH_2O/l/s].

[c] For each of the experiments, the AutoPEEP detection provides: the number of breaths detected as AutoPEEP (denoted as Pos for Positive), the number of breaths detected as NON-AutoPEEP (denoted as Neg for Negative) and the overall label for the considered setting.

Table 5.1 — AutoPEEP detection results provided by the proposed detectors on emulated flow data.

It can be seen that, all the 13 cases were successfully analyzed by the two proposed methods: the Single-breath RDT-based detector and the Sequential RDT-based one. 7 cases were notified with the presence of AutoPEEP and 6 other cases were found normal. Moreover, in each case, all the breaths were precisely classified. Not a single detection error was found among the 323 breaths analyzed. Within the 13 settings considered in this evaluation, the two proposed methods yielded similar detection results for the same reason discussed in Section 5.3.3 below.

5.3.3 Analysis of clinical data

The detection has been validated by simulations and then assessed by emulations with realistic setting. For further evaluation, the AutoPEEP detection framework were

tested *ex-vivo* on various real-world patient curves. A statistics on the detection results could be also meaningful.

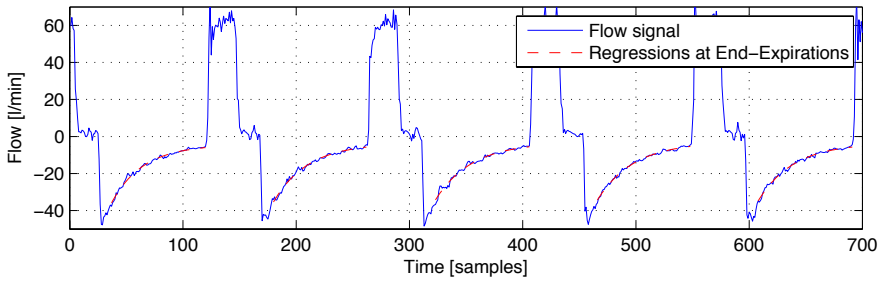
For this purpose, patient flow curves were retrospectively extracted from data files issued from the Medical Intensive Care Unit of Brest University Hospital, France and from the Institut Universitaire de Cardiologie et de Pneumologie de Québec, Canada. These respiratory data files were anonymously recorded on patient undergoing mechanical ventilation. All the flow data were then mixed up to form a unique dataset. In total, the final dataset contains 1998 breaths from 15 patients with different health conditions and following different specific treatments. The parameters of the ventilator also varied depending on the situation and are unknown to the analysis. According to the retrospective aspect of the study and to the fact that the files were anonymized, the study was considered to be in accordance with French legislation by our local ethics committee.

The analysis was performed both manually by a set of experts and automatically by the proposed methods. On the one hand, each breath was carefully screened by two experts of the domain. These experts performed a dual analysis, separately, before confronting their points of view and delivering a final assessment of the data. It is known that a well-trained clinician can easily point out AutoPEEP by visually monitoring the flow curve. Therefore, the decision given by the dual experts could be regarded as the ground-truth label (AutoPEEP/NON-AutoPEEP) for each for each of the breaths they examined. On the other hand, the proposed detectors were used to predict the label of every breath of the dataset. The two analyses were carried out independently and anonymously. The results were then compared together to evaluate the detection performance of the proposed methods.

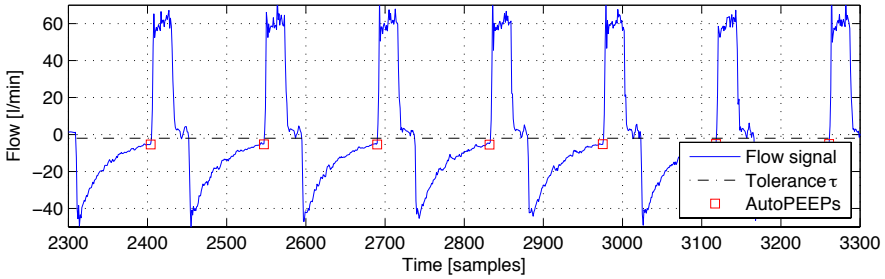
In these experiments, the tolerance was set to $\tau = 2$ [l/min] as before. In this respect, the considered dataset includes 1383 breaths with AutoPEEP and 615 breaths with NON-AutoPEEP. The dataset is somehow unbalanced with the presence of AutoPEEP in 69% of the cases. For the AutoPEEP detectors, level γ was set to 0.01 as usual.

Figure 5.12 presents a typical case with the regression at end-expiration and the corresponding detection. It can be seen that the detection algorithm can precisely reveal the true label for all the breaths. In this example, AutoPEEP is found. Another example with negative decision (i.e. NON-AutoPEEP) is given in Figure 5.13.

To quantitatively assess the detection performance of the proposed methods, we considered four usual evaluation measures: Accuracy, Precision, Recall (Sensitivity) and Specificity. These measures are defined as in Section 2.3.3, where: the number of true positives *TP* (resp. the number of true negatives *TN*) is defined as the number of breaths with (resp. without) AutoPEEP that are correctly predicted; *FP* (false positive) is the number of breaths without AutoPEEP that are falsely predicted as AutoPEEP, and *FN* (false negative) is the number of breath with AutoPEEP that

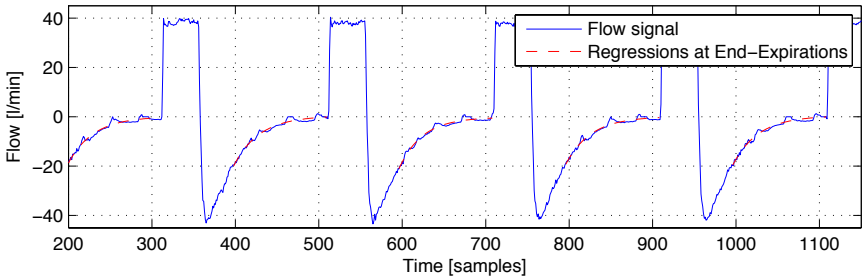


(a) Regressions at End-Expirations

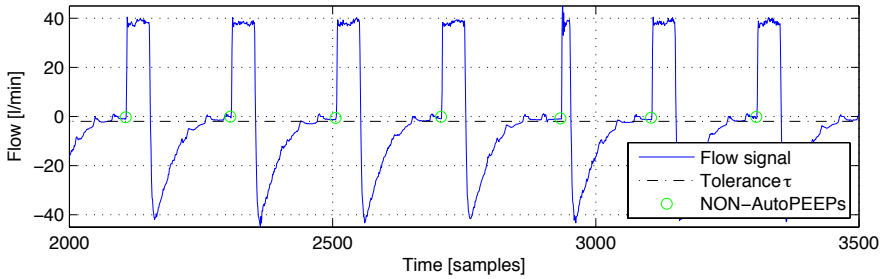


(b) AutoPEEP Detection

Figure 5.12 — Detection results on clinical data, a case with positive decision (i.e. with the presence of AutoPEEP).



(a) Regressions at End-Expirations



(b) AutoPEEP Detection

Figure 5.13 — Detection results on clinical data, a case with negative decision (i.e. NON-AutoPEEP).

are not detected. These four values *TP*, *FP*, *TN*, and *FN* form the so-call confusion matrix of the detection. In terms of the four aforementioned evaluation measures, the

performance results for the two proposed detectors are reported in Table 5.2.

Measure	Single-breath RDT-based detector	Sequential RDT-based detector
Accuracy	**93.09%**	**93.09%**
Precision	99.44%	99.37%
Recall	**90.53%**	**90.60%**
Specificity	98.86%	98.70%

The experiments were carried out with $\tau = 2$[l/min]. For both the detectors, the level was set to $\gamma = 0.01$, which corresponds to an average of 1 false-alarm per 5 minutes (with the usual breathing frequency of 20 [breaths/min]).

Table 5.2 — Detection performance with flow data from patients.

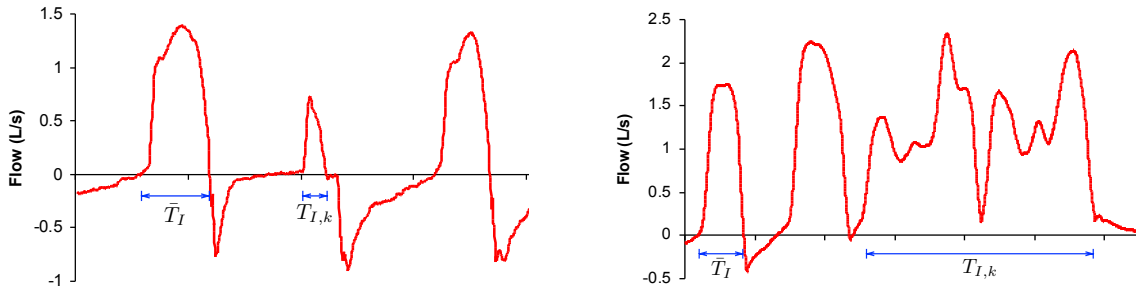
The results show that both the detectors worked very well on patient data with an accuracy higher than 93%, a precision higher than 99%, a recall (sensitivity) higher than 90% and a specificity higher than 98%. For the considered dataset, the two proposed AutoPEEP detectors provided similar results. It is worth mentioning that, by reducing the noise impact, the Sequential RDT-based detector is aimed at improving the detection performance of the Single-breath detector in case the latter fails to reveal the so-called *'twilight region'* AutoPEEP, i.e. AutoPEEP with an end-expiration flow value near the given tolerance τ. Thence, the higher the number of twilight-region AutoPEEPs in the dataset, the more significant the performance improvement can be observed. However, in the considered clinical dataset, the number of twilight region AutoPEEPs, which also presents difficulty to the clinician's eyes in analysis, was very limited. Therefore, no significant difference in detection performance could be seen. However, the use of the Sequential RDT-based detector is recommended for better performance and robustness.

5.4 Extension to detection of asynchrony

Several possible extensions of the detection framework can be suggested for the design of a more complete system. For instance, although being proposed for AutoPEEP detection, the framework also provides elements for the detection of patient-ventilator asynchrony. In this section, although not intended to be an exhaustive study, some types of asynchrony are taken into account. On the basis of how they can be observed and be detected, let us classify patient-ventilator interaction asynchrony into two different categories: one caused by imperfect trigger timing and the other related to the waveform distortion of the observed signal. In what follows, these two categories are investigated with particular examples: for the former one, short cycles, prolonged inspirations and double triggering will be considered; for the latter one, ineffective effort during expiration is taken into account.

5.4.1 Trigger timing related asynchrony

As aforementioned, for this category, the detection of short cycles, prolonged inspirations and double triggering is tackled. In Figure 5.14(a) and Figure 5.14(b), examples of patient flow signals with the presence of, respectively, short cycle and prolonged inspiration are shown. Basically, short cycle and prolonged inspiration concern the amount



(a) Short cycle (b) Prolonged inspiration

Figure 5.14 — Examples of flow signal with short cycle and prolonged inspiration from patient data

of time given to the inspiratory phase of a breath. When this is too short — more precisely,

$$T_{I,k} < \frac{1}{2}\,\bar{T}_I$$

where $T_{I,k}$ is inspiratory time of the k-th breath and $\bar{T}_I$ is the reference value —, a short cycle is said to have occurred. Similarly, an inspiration is said prolonged when the inspiratory time is too long, such that:

$$T_{I,k} > 2\,\bar{T}_I.$$

The reference inspiratory time $\bar{T}_I$ is defined by averaging over previous breaths without timing asynchrony. Empirically, a number of 5 normal breaths are enough to compute this value in practice. On the other hand, double triggering refers to cases where two ventilatory cycles are triggered by the mechanical ventilator within a single patient effort. On the flow curve, double triggering can be revealed by the absence or nearly absence — i.e. presence with a very short duration — of an expiratory phase. In Figure 5.15, an example of the flow signal with double triggering is displayed.

With respect to their definitions, asynchronies of this category can be detected by determining the respiratory phase changes, including inspiration start, inspiration end, expiration start, expiration end, based exclusively on the available flow signal. The detection thus resorts to Problem 2 **[Chg.]** – *Change point detection* formulated in Chapter 4 and the proposed solution is introduced in Section 4.2. In the automatic monitoring platform, the detection of these types of asynchrony is simply a direct

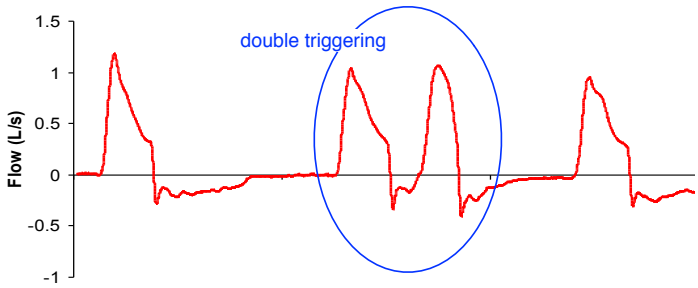


Figure 5.15 — An example of patient flow signal with double triggering

extension, since the required critical instants can be obtained from the phase change detector presented in Section 5.2.3.

5.4.2 Waveform related asynchrony

In this category, asynchrony can be regarded as the deformation of the waveform — or in other words, the distortion of the observation — from some reference curve. This reference can be specified by clinician or can be estimated via previous normal cycles (breaths). The detection of these asynchronies is then Problem 3 **[Dis.]**– *Detection of signal distortion in a time interval* mentioned in Chapter 4 with the solution within the RDT framework detailed in Section 4.3. As a typical example, the detection of ineffective effort during expiration (IEE), a frequent patient-ventilator asynchrony during mechanical ventilation, is hereafter investigated.

It should be noted that ineffective effort (or ineffective triggering) occurs when the mechanical ventilator fails to response to inspiratory effort of patient undergoing ventilatory support. Different factors might relate to this type of interaction failure, including: insufficient inspiratory effort and the presence of AutoPEEP, which increases the necessary effort for triggering [Mulqueeny et al., 2007]. Similar to AutoPEEP, ineffective effort can be visually detected by clinician at the patient bedside via physiologic curves available on ventilator. Various methods for automatizing this detection have also been proposed (see [Blanch et al., 2012, Mulqueeny et al., 2007, Younes et al., 2007] amongst others). Ideally, the combination of the usual respiratory waveforms with the esophageal pressure curve, which captures the patient's inspiratory muscle activity, will provide golden information to analyze ineffective effort. However, such invasive technique is impractical for a monitoring task in real-world application. Fortunately, based exclusively on the flow signal, ineffective effort during expiration (IEE) can be observed. Indeed, it exhibits a deformation of the flow waveform from the usual exponential one, as can be seen in Figure 5.16. In what follows, the analysis of this kind of deformation will then be casted into the RDT framework to carry out the detection of IEE.

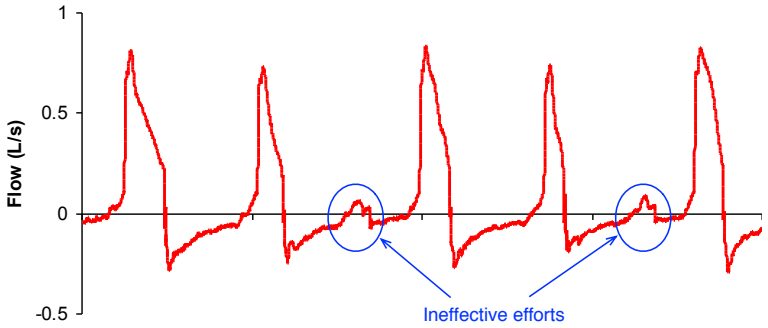


Figure 5.16 — An example of patient flow signal with ineffective efforts during expiratory phases.

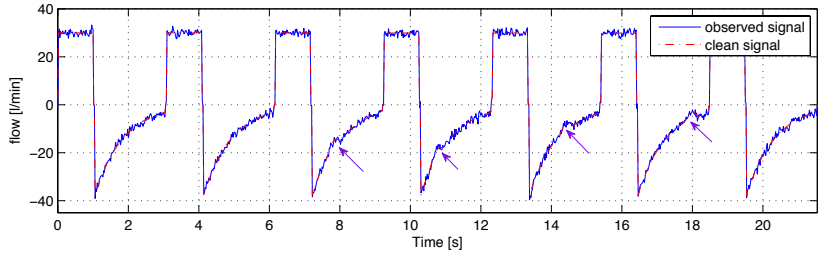
Let f_t be the clean signal and y_t be its observation in additive gaussian noise as in Section 5.2. Prior to the detection of IEE, the expiratory phase of each considered breaths must be identified and well segmented. This processing can easily be achieved by the phase change detector proposed in Section 5.2.3. Let t_k be the end-expiration and L_E be the number of flow signal samples in expiratory phase of the k-th breath. In vector form, the observation model is then $\mathbf{Y}_k = \mathbf{f}_k + \mathbf{X}_k$, where $\mathbf{Y}_k = [y_{t_k-L_E+1}, ..., y_{t_k-1}, y_{t_k}]^T$ (resp. $\mathbf{f}_k = [f_{t_k-L_E+1}, ..., f_{t_k-1}, f_{t_k}]^T$) is the observation (resp. clean flow signal) vector and $\mathbf{X} \sim \mathcal{N}(0, \mathrm{C})$ is gaussian noise with zero mean and covariance matrix C. Let $\mathbf{f}_0$ be the reference vector, which presents the distortion-free flow signal when no IEE is present. Given tolerance τ specified by clinician, the IEE detection then amounts to carrying out the event testing with $[h_0 : \|\mathbf{f}_k - \mathbf{f}_0\| \leqslant \tau]$ (i.e. there is not IEE) and $[h_1 : \|\mathbf{f}_k - \mathbf{f}_0\| > \tau]$ (i.e. there is IEE). The problem is RDT and, with regard to in Section 5.2.3, given a level γ specified by clinician, the decision rule is as follows:

$$[d_{\mathrm{IE}}(\mathbf{Y}_k)] = \begin{cases} \textbf{if} \quad \|\mathbf{Y}_k - \mathbf{f}_0\| > \lambda_\gamma(\tau) \quad \textbf{then} \quad \text{IEE is present } (h_1) \\ \textbf{if} \quad \|\mathbf{Y}_k - \mathbf{f}_0\| \leqslant \lambda_\gamma(\tau) \quad \textbf{then} \quad \text{IEE is not found } (h_0) \end{cases} \tag{5.9}$$

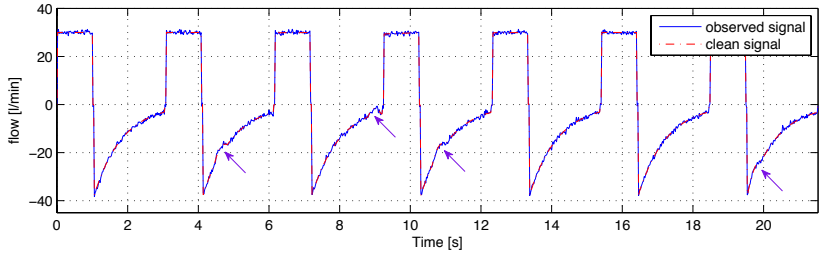
In practice, either $\mathbf{f}_0$ is known or it can be estimated from previous normal cycles that present no distortion. Since flow signal during expiratory phase exhibits a specific exponential waveform, the regression curve via a robust method can also be considered as a reference.

As a preliminary detection performance assessment, simulations were carried out. Flow signal was first synthesized on computer. The signal was generated with a breath rate equal to 20 [breaths/min], an inspiratory-to-expiratory time ratio $I{:}E = 1:2$ and the sampling time $T_s = 0.02$ [s]. As a result, the dimension of the problem was then $L_E = 100$. The centered gaussian noise was supposed independent and identically distributed with variance σ^2, i.e. $\mathrm{C} = \sigma^2 \mathbb{1}_{L_E}$. The presence/absence of IEE in a breath was random with equal probability. The duration of simulated patient effort was set to $T_{\mathrm{es}} = 0.4$ [s]. Its position (when present) was uniformly distributed along the expi-

ratory phase and its amplitude was rather small (the maximum value of P_{mus} is -0.5 [cmH_2O]). In Figure 5.17, examples of flow signal with different observation noise level were presented. The presence of IEEs was also highlighted.



(a) $\sigma = 1$ [l/min] (i.e. $\frac{\tau_0}{\sigma} = -20$dB)



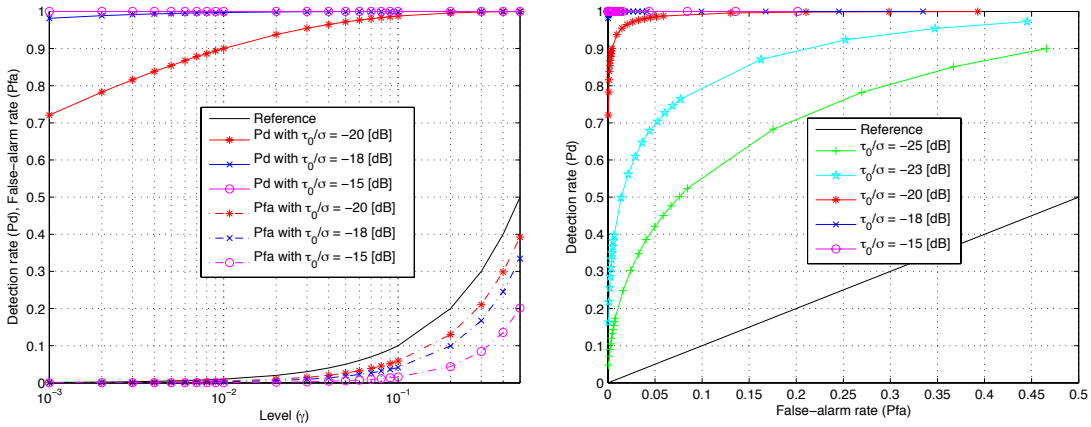
(b) $\sigma = 0.5623$ [l/min] (i.e. $\frac{\tau_0}{\sigma} = -15$dB)

Figure 5.17 — Example of synthesized flow signal with different noise levels. The presence of IEE is pointed out by small arrows. It can also be seen that the patient effort was set to be rather small in comparison to the noise level.

For the detection, the estimated flow signal yielded by robust regression with respect to the model function (5.6) is used as the reference. Observation from the first normal breaths were used for such estimation. The tolerance was set to $\tau = \frac{T_{\text{es}}}{T_s}\tau_0 = 20\tau_0$ with $\tau_0 = 0.1$ [l/min]. As a result, $\tau = 2$ [l/min]. Indeed, $\frac{T_{\text{es}}}{T_s} = 20$ is merely the expected number of distorted samples in the observation vector. In practice, this number can also be estimated by counting the number of significant wavelet coefficients of the difference vector $\mathbf{Y}_k - \mathbf{f}_0$.

The detection performance is reported in Figure 5.18. In Figure 5.18(a), the curves representing the false-alarm and detection probabilities with respect to different values of level γ are given. The results show that, even being masked by rather strong observation noise, IEE can successfully be revealed with high precision. The false-alarm rate is always guaranteed to be lower than the specified level γ. The Receiver Operating Curves (ROCs) are also depicted in Figure 5.18(b) for reference. The simulation results suggest that the proposed RDT-based IEE detector provides high detection performance. However, it is required that further assessment is carried out on clinical data from patients.

It should be noted that, by using the RDT framework, no prior distortion motifs are preferred. Therefore, any type of patient-ventilator asynchrony that exhibits a



(a) P_d and P_{fa} (b) ROCs

Figure 5.18 — IEE detection performance. On the left (i.e. Figure 5.18(a)), the false-alarm probability and the detection probability were plotted with respect to different values of γ. On the right (i.e. Figure 5.18(b)), although being less meaningful in RDT framework, the Receiver Operating Curves (ROCs) with different observation noise level were provided.

distortion in the flow signal with respect to the reference is considered in the same way and can also be detected. In that case, the aforementioned method can serve as a first-stage detector to notify the presence of asynchrony then a second-stage classifier will be employed to identify what type the detected asynchrony might be.

In this section, only the flow signal gets involved since it is available on most of currently used mechanical ventilators. It is, however, expected that, when they are available, others respiratory signals (such as: pressure, volume, ...) can be taken into account to yield better detection results. In this respect, it is also worth mentioning that some types of asynchrony (for examples, initial overshoot and final overshoot) cannot be revealed based exclusively on flow signal. In such cases, other signals are needed.

5.5 Discussions

5.5.1 Automatic detection of ventilatory support failure

The experiment results of Section 5.3 have shown that, based exclusively on the flow signal available on most of currently used mechanical ventilators, the proposed detection framework is capable of precisely detecting AutoPEEP — one of the most frequent patient-ventilator interaction abnormalities. To the best of our knowledge, this is the first work on the automatic detection of AutoPEEP for continuous monitoring the patient-ventilator interface during mechanical ventilation. The RDT and its extension

in sequential framework have been successfully applied to provide high performance detectors of AutoPEEP.

The proposed detection algorithms have very low complexity and require very little computational power. The proposed detection framework can then be deployed as a real-time functional block, which, in turn, can either be integrated into most of currently used ventilator to carry out the monitoring as a novel additional functionality or be developed as an independent device to control the functioning of mechanical ventilators in use.

For evaluation on real-world data, retrospective data files have been used in this study. With a double-blinded and dual expert analysis on these data, we were already able to assess whether the decisions given by the proposed framework were in accordance with those judged by the experts. In the next validation step, continuous and prospective recordings of respiratory signals will be performed to get better insight into cases where any disagreement between the proposed system and the therapist might occur. Furthermore, it is also meaningful to perform a semi-closed-loop analysis, in which the therapist supervises, validates the decisions yielded by the proposed detection system and adjusts the corresponding ventilatory parameters to correct any possible abnormality.

The focus of this study is on the detection of ventilatory support failure, for which, only the presence/absence of an abnormality is considered. It is however expected that the estimation, which provides a quantitative evaluation of the identified of normality, might also be of interest. This kind of information will also be helpful for the clinician to judge the situation and to effectively tune the parameters. For further study, possible corrections with respect to each type of abnormality could also be suggested.

5.5.2 Real-time remote monitoring framework

With the introduction of the automatic detection of ventilatory support failure in a continuous real-time monitoring, the clinicians can optimize their tasks and thus improve care. This is of interest since clinicians cannot always be near the patient's bedside. The constraint becomes more crucial for patients treated by home mechanical ventilation. It should be noted that with a huge number of ventilators used at home (more than 600 000 in France and more than 1 million in the United States of America), such automatic monitoring with the supervision of a clinician from distance is indeed necessary. In this respect, a more complex framework supporting real-time remote monitoring is proposed. The whole solution can be summarized as in Figure 5.19.

The framework can be divided into two ends linked together via a central server. On one end, each mechanical ventilator is equipped with a real-time automatic monitoring module capable of communicating with the server via an available physical medium of

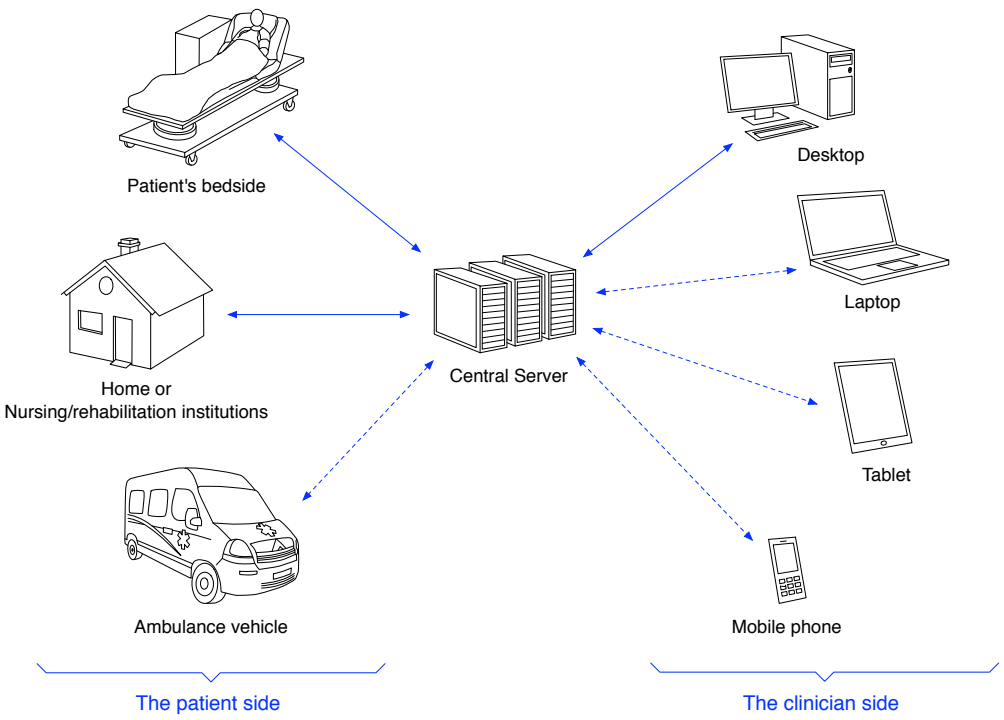


Figure 5.19 — Realtime remote monitoring solution for ventilatory support failure detection

any type. These ventilators can be installed at hospitalized patients' bedside, at home, in nursing/rehabilitation institution or even on ambulance vehicles, as depicted in Figure 5.19. The communication links from these ventilators to the central server can be implemented with any state-of-the-art technology, such as a wired/wireless Internet connection, a communication link via satellite, etc. On the other end, communication devices are available for clinician to perform their supervisory tasks. These devices can be their desktops connected to the server via an Ethernet cable, their laptops/tablet with a WiFi links or even their mobile phones with standard cellular network connection. The central server manages all the data exchanges among different equipments in the framework. It can receive data from monitoring modules, processes the observed signals and forward necessary information to clinicians. Depending on the throughput of the transmission link, the automatic monitoring module can either process data and send only the detection result or transfer the whole respiratory signals to the server, or both. Similarly, the central server can either simply notify clinicians when an abnormality is found or send them the complete respiratory data and automatic analysis result for their supervision. The clinicians' communication devices must support data visualization for clinical analysis.

The advantage of such a monitoring solution is that, not only it can help alleviate therapists' amount of work and improve care on patients, especially those treated at home, but it can also get multiple clinicians involved in supervising a case, regardless

the possible distance. For example, the clinician in an ambulance can easily consult an expert of domain, who is not present in the vehicle, and send respiratory data in real-time for his/her analysis. In another aspect, the framework makes it easy to record and collect data to provide a database for further works, in which machine learning based methods can play an important role.

5.5.3 Virtual ventilatory support simulator

It must be emphasized that evaluation on clinical data recorded from real-world patients is the *golden standard* for assessing the proposed methods. Such data, however, represent some unavoidable limitations. Firstly, since data can only be recorded on patients enrolled in the study, it might happen that data on some category of patient in unavailable. Therefore, an exhaustive study on patients with various health conditions can hardly be carried out. Moreover, even for a specific category, the number of enrolled patients may not be enough for a statistical analysis to be meaningful. Secondly, such study on patients — requiring a permission from an ethnic committee — might become laborious and, therefore, the amount of data might be limited. Finally, a closed-loop test on real-world patient might strictly be regulated for safety sake and hardly be done because of possible harmfulness it might yield to patient. In these respects, a virtual ventilatory support simulator as shown in Figure 5.20 is developed.

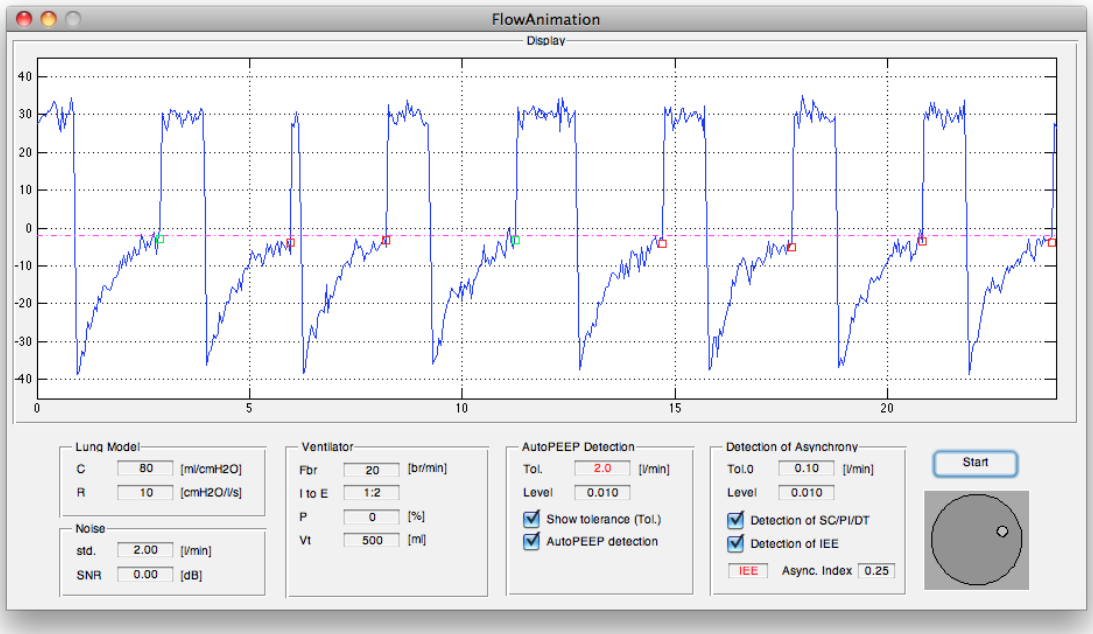


Figure 5.20 — The GUI (Graphical User Interface) of the implemented virtual ventilatory support simulator

This virtual simulator can simulate both the behavior a human respiratory system and the action of a mechanical ventilator. On the patient side, different mechanical characteristics of the respiratory system can be easily parametrized, making it possible to mimic various categories of patients in practice. On the mechanical ventilator side, the control parameters are similar to those in standard ventilators currently used in practice. Other environment parameters are also adjustable, including noise. More details are given in Appendix D. Such a virtual simulator facilitates the assessment on various configurations, involving different types of patients and different ventilator settings, including closed-loop test. New algorithms can easily be integrated for different detection, estimation, correction tasks.

In the first evaluation phase of the proposed real-time remote monitoring framework, virtual simulators can be used instead of real mechanical ventilators. This replacement not only reduces the cost of the test but also brings flexibility to perform assessment in complex situations.

Moreover, the proposed virtual environment is very helpful for educational purposes. Clinicians could be trained with situations that are rarely present in practice. They can virtually manipulate their corrections on ventilatory parameters and visualize the patient's response. It could also be noted that such training course on this virtual environment costs less than classical ones, relying on real mechanical ventilators and respiratory system analog. The model can be easily replicated and employed in any educational institutes.

Conclusion

In comparison with either parametric methods, which offer optimal properties, or non-parametric approaches, which bring robustness, the RDT framework combines the benefits of both the parametric and non-parametric ones. On the one hand, the signal is supposed to be random with unknown distribution. None of its statistical characteristics is taken into account. The suggested tests can thus guarantee some robustness against signal distribution variations. On the other hand, the RDT framework provides statistical optimality in a sense similar to Wald's. More specifically, the power of the tests is optimized while the probability of false-alarm is still upper-bounded by some specific level. In this respect, to some extent, RDT could be considered as a semi-parametric approach. Unlike machine learning based methods, to compensate the lack of prior knowledge on the conditional distribution of observations under the permissible hypotheses, the mere information needed to perform RDT concerns solely the noise covariance matrix, which can often be estimated in practice. No training phase, which implies the use of a training dataset, is required. Therefore, RDT is suitable for many applications in which collecting data and annotating a sufficiently large and representative dataset concerning signal is a laborious task. Furthermore, by introducing the notion of tolerance, RDT makes it possible to incorporate the experience on the domain into the decision. The calculation of optimal thresholds for testing is based not only on the nature of the observation but also on the tolerance value selected by experts and the false-alarm level specified on the basis of practical requirements. This makes the RDT framework flexible and, therefore, easier to be adapted to various applications.

The RDT framework is very general and can be applied in many domains. In Sections 3.3, 4.1, 4.2, 4.3, its applications to classical signal processing problems have been presented. In signal detection, the tests suggested by RDT have been shown to be resistant to random signal model mismatch, for which standard Neyman-Pearson's likelihood test might fail. Therefore, for certain cases where it is difficult to characterize precisely the difference between the way the testing problem is formulated and what the actual observation might be, RDT could be an alternative to conventional likelihood theory. On the other hand, three typical signal deviation/distortion detection problems have been formulated and cast into the RDT framework. As illustrating ex-

amples, the solutions to these three problems have been successfully applied to different tasks in a continuous monitoring framework for automatic detection of AutoPEEP and asynchrony during mechanical ventilation. Beyond the biomedical engineering domain, other applications, in which a deviation/distortion from a nominal reference must be detected but can hardly be modeled, could also be investigated, such as: tracking, anti-collision radar, robotics, structural health monitoring, fault-detection, etc.

As an extension of the RDT framework, the Sequential RDT has been proposed with the introduction of the optimal dual-threshold thresholding tests. In automatic AutoPEEP detection, the resulting sequential detector has been shown to yield better performance than the single-breath RDT-based one. The detection rate is improved while the specified bound on the false-alarm rate is still respected. The success of the Sequential RDT suggests its use in practice when more available observations can be profited to provide more evidence to the decision. These observations could be either temporally or spatially distributed. For example, in a sensor network, when observation from a particular sensor does not provide enough information for a confident decision, either more samples from that sensor could be requested or observations from other ones could be consulted.

General conclusion and perspectives

In this thesis, two trends in non-parametric statistics have been considered. On the one hand, the machine learning has been considered with its application to the detection of protein interface hotspot. A new family of hotspot descriptors, which might provide better insight into the nature of protein interaction and protein hotspots, has also been proposed. On the other hand, a robust hypothesis testing framework, namely RDT, has been presented with its application in automatic detection of AutoPEEP and asynchrony during mechanical ventilation. Classical signal processing problems such as signal detection and the three typical signal deviation/distortion detection have also been successfully cast into such a framework. Moreover, an extension of the RDT framework in sequential analysis has also been developed.

The machine learning imposes no prior hypothesis on the observation distributions. Therefore, it can deal with any kind of descriptors. This advantage makes machine learning suitable for various applications in which the nature of features (observations) and how they relate to the decision output are hardly known. However, its requirement for a well-descriptive training dataset presents difficulty in many cases and no guarantees on the performance can be assured. Additionally, although the prediction performance of a machine learning based method might be high, it is hard to interpret how the ensemble of descriptors determines the true output. Nevertheless, descriptors can be ranked and those relevant to the problem of interest can be selected for any ambitious study on this complicated relation.

Although it is non-parametric, RDT requires a prior observation model and the noise covariance to be known. This however limits its use in some applications. In contrast to machine learning, RDT guarantees a maximum value for the false-alarm rate and provides certain optimality with respect to the power of the test.

In this study, an extension of RDT in sequential detection framework has been shown to be successful for the detection of unidimensional signal deviations at critical instants. In a future work, a general approach with d-dimensional signal in Sequential

RDT could be investigated. Some other theoretical aspects, such as: delay to decision, confidence interval and other statistical optimalities of a sequential test, should also be given.

For the moment, RDT considers the observation in additive noise with normal distribution. It can be extended to other types of noises, such as Generalized Gaussian or Gaussian mixture distributed ones. A further study of this type can give a significant impact in practice.

The combination of RDT with Machine learning could also be a potential trend. For instance, since the split at each node in a decision tree involves a thresholding on the descriptor, it is expected that RDT could be profitable to yield better results. Similarly, other tasks which relies on the choice of an optimal threshold could also be improved within RDT.

As mentioned in Chapter 2, the detection of protein interaction hotspots can be considered as the location of high-energy regions in certain transform domains. Therefore, the thresholding test provided in the RDT framework could be investigated to carry out such a detection task. On the other hand, in mechanical ventilation monitoring application, as long as some further asynchrony classification task is required, a machine learning scheme can be used to analyze different distortion motifs in the observed respiratory signals.

APPENDIX A

Constraint violation of Neyman-Pearson likekihood test under model mismatch

Given observation under the model mismatch as in (3.27), if the Neyman-Pearson's likelihood test $\mathcal{T}_{\text{NP}}$ is still applied, the probability of false-alarm is expressed as:

$$\begin{aligned} \text{P}_{\text{fa}}[\mathcal{T}_{\text{NP}}] &\overset{\text{def}}{=} \text{P}\left[\boldsymbol{\xi}_1^T \text{C}^{-1}(\boldsymbol{\Xi}_0 + \mathbf{X}) > \lambda_{\text{NP}}\right] \\ &= \text{P}\left[\boldsymbol{\xi}_1^T \text{C}^{-1}\mathbf{X} > (\lambda_{\text{NP}} - \boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\Xi}_0)\right] \\ &= \text{P}\left[\frac{\boldsymbol{\xi}_1^T \text{C}^{-1}\mathbf{X}}{\sqrt{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\xi}_1}} > \frac{\lambda_{\text{NP}} - \boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\Xi}_0}{\sqrt{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\xi}_1}}\right] \end{aligned} \tag{A.1}$$

Now put $\bar{X} = \frac{\boldsymbol{\xi}_1^T \text{C}^{-1}\mathbf{X}}{\sqrt{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\xi}_1}}$, $\bar{\lambda}_{\text{NP}} = \frac{\lambda_{\text{NP}}}{\sqrt{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\xi}_1}}$ and $\bar{\Xi} = \frac{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\Xi}_0}{\sqrt{\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\xi}_1}}$. It should be noted that $\bar{X} \sim \mathcal{N}(0,1)$ since $\mathbf{X} \sim \mathcal{N}(0, \text{C})$ and that $\bar{\Xi}$ possesses a symmetric probability density function $f_{\bar{\Xi}}(\xi)$ since $\boldsymbol{\xi}_1^T \text{C}^{-1}\boldsymbol{\Xi}_0$ is supposed to be symmetrically distributed. Furthermore, $\bar{\Xi}$ and $\bar{X}$ are independent. Equation (A.1) can therefore be rewritten as:

$$\begin{aligned} \text{P}_{\text{fa}}[\mathcal{T}_{\text{NP}}] &= \text{P}\left[\bar{X} > \bar{\lambda}_{\text{NP}} - \bar{\Xi}\right] \\ &= 1 - \int_{-\infty}^{+\infty} \Phi(\bar{\lambda}_{\text{NP}} - \xi) f_{\bar{\Xi}}(\xi) \text{d}\xi \\ &= 1 - \int_{0}^{+\infty} [\Phi(\bar{\lambda}_{\text{NP}} - \xi) + \Phi(\bar{\lambda}_{\text{NP}} + \xi)] f_{\bar{\Xi}}(\xi) \text{d}\xi \end{aligned} \tag{A.2}$$

where $\Phi(.)$ is the cumulative distribution function of any random variable that follows the standard normal distribution. We also have $\gamma = \text{P}_{\text{fa}}^*[\mathcal{T}_{\text{NP}}] \overset{\text{def}}{=} \text{P}\left[\boldsymbol{\xi}_1^T \text{C}^{-1}\mathbf{X} > \lambda_{\text{NP}}\right] = 1 - \Phi(\bar{\lambda}_{\text{NP}})$, therefore, $\bar{\lambda}_{\text{NP}} > 0$ for any $\gamma < 0.5$.

Additionally, it could be proved that

$$\Phi(\bar{\lambda}_{\text{NP}} - \xi) + \Phi(\bar{\lambda}_{\text{NP}} + \xi) < 2\Phi(\bar{\lambda}_{\text{NP}}) \tag{A.3}$$

for any $\bar{\lambda}_{\text{NP}} > 0$ and $\xi > 0$. Indeed, this really holds when $\xi \leqslant \bar{\lambda}_{\text{NP}}$ because the function $\Phi(x)$ is strictly concave (or concave downwards) for $x \geqslant 0$. When $\xi > \bar{\lambda}_{\text{NP}}$,

the inequality can be proved as follows. Let $\phi(.)$ be the probability density function (pdf) of any random variable following standard normal distribution. Since $\phi(.)$ is strictly decreasing in $[0, +\infty)$, the function

$$g(a) = \Phi(a + \Delta) - \Phi(a) = \int_{a}^{a+\Delta} \phi(x)\mathrm{d}x$$

is also strictly decreasing in $[0, +\infty)$ for any $\Delta > 0$. Therefore, on the one hand,

$$\Phi(\bar{\lambda}_{\mathrm{NP}}) - \Phi(0) > \Phi(2\bar{\lambda}_{\mathrm{NP}}) - \Phi(\bar{\lambda}_{\mathrm{NP}}) \tag{A.4}$$

and, on the other hand,

$$\Phi(0) - \Phi(\bar{\lambda}_{\mathrm{NP}} - \xi) = \Phi(\xi - \bar{\lambda}_{\mathrm{NP}}) - \Phi(0) > \Phi(\xi + \bar{\lambda}_{\mathrm{NP}}) - \Phi(2\bar{\lambda}_{\mathrm{NP}}). \tag{A.5}$$

The result then follows by adding inequations (A.4) and (A.5) side-by-side:

$$\Phi(\bar{\lambda}_{\mathrm{NP}} - \xi) + \Phi(\bar{\lambda}_{\mathrm{NP}} + \xi) < 2\Phi(\bar{\lambda}_{\mathrm{NP}}). \tag{A.6}$$

Now, since $\bar{\lambda}_{\mathrm{NP}} > 0$ and ξ varies in $(0; +\infty)$, using Inequality (A.3) to upper bound the integrand in the right hand side of (A.2), we have:

$$\mathrm{P}_{\mathrm{fa}}[\mathcal{T}_{\mathrm{NP}}] > 1 - 2\Phi(\bar{\lambda}_{\mathrm{NP}}) \int_{0}^{+\infty} f_{\bar{\Xi}}(\xi)\mathrm{d}\xi \tag{A.7}$$

and thus

$$\mathrm{P}_{\mathrm{fa}}[\mathcal{T}_{\mathrm{NP}}] > 1 - \Phi(\bar{\lambda}_{\mathrm{NP}}) = \gamma \tag{A.8}$$

The Neyman-Pearson's constraint on the probability of false-alarm is then violated.

APPENDIX B

The convergence of the two thresholds in the proposed dual-threshold test

To begin with, remind that $\lambda_\gamma(\rho)$ (resp. $\lambda_{1-\gamma}(\rho)$) is the unique solution in η to the equation $1-\gamma = [\Phi(\eta-\rho) - \Phi(-\eta-\rho)]$ (resp. $\gamma = [\Phi(\eta-\rho) - \Phi(-\eta-\rho)]$), where $\Phi(.)$ is the c.d.f. of any random variable following the standard normal distribution. Therefore, the two thresholds:

$$\begin{array}{l} \lambda_{1:K}^{(h)} = \sigma_{W,K}\lambda_\gamma\left(\frac{\tau}{\sigma_{W,K}}\right) \\ \lambda_{1:K}^{(\ell)} = \sigma_{W,K}\lambda_{1-\gamma}\left(\frac{\tau}{\sigma_{W,K}}\right) \end{array} \tag{B.1}$$

of the proposed dual-threshold test are respectively the unique solutions in η to:

$$\begin{array}{rcl} 1-\gamma & = & F_\tau(\eta, \sigma_{W,K}) \\ \gamma & = & F_\tau(\eta, \sigma_{W,K}) \end{array} \tag{B.2}$$

where $F_\tau(\eta,\sigma) = \Phi\left(\frac{\eta-\tau}{\sigma}\right) - \Phi\left(\frac{-\eta-\tau}{\sigma}\right)$. Let us now consider the variation of function $F_\tau(\eta,\sigma)$ with respect to (w.r.t) its variables η and σ.

Denote by $\phi(.)$ the pdf of any standard normally distributed random variable. On the one hand, we have:

$$\frac{\partial F_\tau(\eta,\sigma)}{\partial \eta} = \frac{1}{\sigma}\left(\phi\left(\frac{\eta-\tau}{\sigma}\right) + \phi\left(\frac{-\eta-\tau}{\sigma}\right)\right) > 0 \tag{B.3}$$

Function $F_\tau(.,\sigma)$ is thus strictly *increasing* for all $\sigma > 0$.

On the other hand,

$$\frac{\partial F_\tau(\eta,\sigma)}{\partial \sigma} = -\frac{1}{\sigma^2}\left[\phi\left(\frac{\eta-\tau}{\sigma}\right)(\eta-\tau) + \phi\left(\frac{-\eta-\tau}{\sigma}\right)(\eta+\tau)\right] \tag{B.4}$$

(a) When $\eta \geqslant \tau \geqslant 0$, we have $\frac{\partial F_\tau(\eta,\sigma)}{\partial \sigma} < 0$. Therefore, function $F_\tau(\eta, .)$ is strictly *decreasing*.

(b) When $\tau > \eta \geqslant 0$, we have:

$$\frac{\phi\left(\frac{\eta-\tau}{\sigma}\right)}{\phi\left(\frac{-\eta-\tau}{\sigma}\right)} = \exp\left(\frac{2\eta\tau}{\sigma^2}\right) \xrightarrow{\sigma\to 0} +\infty$$

Therefore, $\frac{\partial F_\tau(\eta,\sigma)}{\partial \sigma} > 0$ when σ is small enough. Function $F_\tau(\eta, .)$ is thus strictly *increasing* w.r.t σ in a sufficiently small interval near 0.

For any γ such that $0 < \gamma < 0.5$, we have $\lambda^{(h)}_{1:K} > \lambda^{(\ell)}_{1:K}$

For $0 < \gamma < 0.5$, we have:

$$1 - \gamma > \gamma$$

Therefore,

$$F_\tau(\lambda^{(h)}_{1:K}, \sigma_{W,K}) > F_\tau(\lambda^{(\ell)}_{1:K}, \sigma_{W,K})$$

Since $F_\tau(\eta, \sigma_{W,K})$ is strictly increasing w.r.t η, it can be inferred that:

$$\lambda^{(h)}_{1:K} > \lambda^{(\ell)}_{1:K}$$

For $0 < \gamma < 0.5$, threshold $\lambda^{(h)}_{1:K}$ is decreasing and tends to tolerance τ when $\sigma_{W,K}$ tends to 0

We have:

$$F_\tau(\tau, \sigma_{W,K}) = 0.5 - \Phi\left(\frac{-2\tau}{\sigma_{W,K}}\right)$$

Since $0 < \gamma < 0.5$, it can be seen that:

$$1 - \gamma > 0.5 > 0.5 - \Phi\left(-\frac{2\tau}{\sigma_{W,K}}\right)$$

And, thence,

$$F_\tau(\lambda^{(h)}_{1:K}, \sigma_{W,K}) > F_\tau(\tau, \sigma_{W,K})$$

Since $F_\tau(., \sigma_{W,K})$ is strictly increasing, it can be inferred that $\lambda^{(h)}_{1:K} > \tau$. Therefore, it follows from (a) above that $F_\tau(\lambda^{(h)}_{1:K}, .)$ is strictly decreasing.

Now, it is noted that, $F_\tau(\lambda^{(h)}_{1:K}, \sigma_{W,K}) = 1 - \gamma = \text{const}$. Moreover, $F_\tau(\lambda^{(h)}_{1:K}, \sigma_{W,K})$ is strictly increasing w.r.t $\lambda^{(h)}_{1:K}$ and strictly decreasing w.r.t $\sigma_{W,K}$. Therefore, $\lambda^{(h)}_{1:K}$ must be strictly increasing w.r.t $\sigma_{W,K}$. In other words, $\lambda^{(h)}_{1:K}$ strictly decreases when $\sigma_{W,K}$ decreases. Additionally, $\lambda^{(h)}_{1:K}$ is lower-bounded by τ. The limit for $\lambda^{(h)}_{1:K}$ when $\sigma_{W,K} \to 0$ exists. It can be shown that this limit is the specified tolerance τ.

Indeed, for arbitrarily small value $\varepsilon > 0$, $F_\tau(\tau + \varepsilon, \sigma_{W,K}) = \Phi\left(\frac{\varepsilon}{\sigma_{W,K}}\right) - \Phi\left(\frac{-2\tau-\varepsilon}{\sigma_{W,K}}\right)$ is a continuous function of $\sigma_{W,K}$, tends to 1 when $\sigma_{W,K} \to 0$ and tends to 0 when $\sigma_{W,K} \to +\infty$. Therefore, there exists σ_ε so that: $1 - \gamma = F_\tau(\tau + \varepsilon, \sigma_\varepsilon)$. Since $\lambda_{1:K}^{(h)}$ is strictly increasing w.r.t $\sigma_{W,K}$, we have $\tau < \lambda_{1:K}^{(h)} < \tau + \varepsilon$ for any $\sigma_{W,K}$ such that $0 < \sigma_{W,K} < \sigma_\varepsilon$. In brief, $\lambda_{1:K}^{(h)} \xrightarrow[\sigma_{W,K} \to 0]{} \tau$.

For $0 < \gamma < 0.5$, threshold $\lambda_{1:K}^{(\ell)}$ tends to tolerance τ when $\sigma_{W,K}$ tends to 0

For $\sigma_{W,K}$ small enough, we have $\gamma < 0.5 - \Phi\left(-\frac{2\tau}{\sigma_{W,K}}\right)$ or, in other words, $F_\tau(\lambda_{1:K}^{(\ell)}, \sigma_{W,K}) < F_\tau(\tau, \sigma_{W,K})$. Therefore, $0 < \lambda_{1:K}^{(\ell)} < \tau$. In what follows, we only consider $\sigma_{W,K}$ in a sufficiently small interval $(0, \sigma_M)$ near 0 so that: $\lambda_{1:K}^{(\ell)} < \tau$ and $F_\tau(\lambda_{1:K}^{(\ell)}, .)$ is strictly increasing.

We have $F_\tau(\lambda_{1:K}^{(\ell)}, \sigma) = \gamma = \text{const}$. Moreover, $F_\tau(\lambda_{1:K}^{(\ell)}, \sigma)$ is strictly increasing w.r.t σ and strictly increasing w.r.t $\lambda_{1:K}^{(\ell)}$. It can be inferred that $\lambda_{1:K}^{(\ell)}$ is decreasing w.r.t $\sigma_{W,K}$. In other words, $\lambda_{1:K}^{(\ell)}$ increases when $\sigma_{W,K}$ decreases. Additionally, $\lambda_{1:K}^{(\ell)}$ is upper-bounded by τ. Therefore, there exists a limit for $\lambda_{1:K}^{(\ell)}$ when $\sigma_{W,K} \to 0$. This limit can be proved to be the tolerance τ as follows.

Given any $\varepsilon > 0$, we have $F_\tau\left(\tau - \varepsilon, \sqrt{\varepsilon}\right) = \Phi\left(-\sqrt{\varepsilon}\right) - \Phi\left(-\frac{2\tau}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\right)$. Thereby, $\lim_{\varepsilon \to 0} F_\tau\left(\tau - \varepsilon, \sqrt{\varepsilon}\right) = \frac{1}{2}$ and $\lim_{\varepsilon \to \tau} F_\tau\left(\tau - \varepsilon, \sqrt{\varepsilon}\right) = 0$. Since $F_\tau\left(\tau - \varepsilon, \sqrt{\varepsilon}\right)$ is a continuous function of ε, we derive from the foregoing the existence of some ε such that $F_\tau\left(\tau - \varepsilon, \sqrt{\varepsilon}\right) = \gamma$. Given $\sigma_{W,K}$ such that $0 < \sigma_{W,K} < \inf(\sqrt(\varepsilon), \sigma_M)$, we then have $\tau - \varepsilon < \lambda_{1:K}^{(\ell)} < \tau$ since $\lambda_{1:K}^{(\ell)}$ strictly decreases w.r.t $\sigma_{W,K}$. Therefore, $\lambda_{1:K}^{(\ell)} \xrightarrow[\sigma_{W,K} \to 0]{} \tau$.

APPENDIX C

Gaussianity of the aggregated noise when using the waveform vector

In this appendix, we will prove that the aggregated noise given by:

$$w_k = \frac{\mathbf{p}_k^T \mathbf{X}_k}{\|\mathbf{p}_k\|^2} \tag{C.1}$$

tends to a normally distributed random variable, as long as L is large enough and the noise samples in $\mathbf{X}_k$ are i.i.d (independent and identically distributed). For the sake of notation simplicity, let us discard index k, which originally indicates that the k-th breath is currently considered.

To begin with, let $\mathbf{p} = [p_1, p_2, .., p_L]^T$ and $\mathbf{X} = [x_1, x_2, ..., x_L]^T$. Since X is centered, we have $\mathrm{E}[x_i] = 0, \forall i = 1..L$. Denote by σ the standard deviation for each x_i and by μ^4 the 4-th order moment of x_i for any $i = 1..L$.

Put $Z_i = \frac{p_i}{\|p\|^2} x_i$, we have:

$$\begin{aligned} \mathrm{E}[Z_i] &= \tfrac{p_i}{\|p\|^2}\mathrm{E}[x_i] = 0 \\ \mathrm{var}\,[Z_i] &= \tfrac{p_i^2}{\|p\|^4}\mathrm{E}[x_i^2] = \tfrac{p_i^2}{\|p\|^4}\sigma^2 \\ \mathrm{E}\left[(Z_i - \mathrm{E}[Z_i])^4\right] &= \tfrac{p_i^4}{\|p\|^8}\mathrm{E}[x_i^4] = \tfrac{p_i^4}{\|p\|^8}\mu^4 \end{aligned} \tag{C.2}$$

Let us now consider the Lyapunov's condition for the central limit theorem [Billingsley, 1995, Theorem 27.4, p. 362] with $\delta = 2$. We have:

$$\zeta = \frac{\sum_{i=1}^{L} \mathrm{E}\left[|Z_i - \mathrm{E}[Z_i]|^{2+\delta}\right]}{\left(\sum_{i=1}^{L} \mathrm{E}\left[(Z_i - \mathrm{E}[Z_i])^2\right]\right)^{\delta}} = \frac{\mu^4}{\sigma^4} \frac{\sum_{i=1}^{L} p_i^4}{\left(\sum_{i=1}^{L} p_i^2\right)^2} \tag{C.3}$$

With respect to equation (5.8), we have $p_i \geqslant 1$ for $i = 1..L$. Since the flow signal is bounded because of physic, there also exists a value A so that $|p_i| \leqslant A$ for every $i = 1..L$. We then have:

$$\begin{aligned} \sum_{i=1}^{L} p_i^4 &\leqslant LA^4 \\ \left(\sum_{i=1}^{L} p_i^2\right)^2 &\geqslant L^2 \end{aligned} \tag{C.4}$$

And, therefore,

$$\zeta \leqslant \frac{\mu^4}{\sigma^4}\frac{A^4}{L} \xrightarrow[L\to\infty]{} 0 \tag{C.5}$$

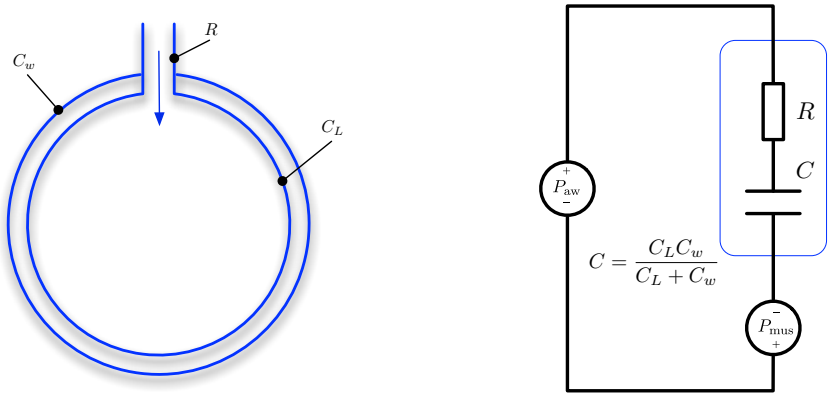
According to the Lyapunov's central limit theorem [Billingsley, 1995, Theorem 27.4, p. 362], $w = \frac{\mathbf{p}^T\mathbf{X}}{\|\mathbf{p}\|^2} = \sum_{0}^{L} Z_i$ converges in distribution to a normal random variable when $L \to +\infty$. The aggregated noise w_k in (5.1) then converges to gaussian noise when L is sufficiently large.

APPENDIX D

Virtual ventilatory support simulator

Principles

It should be noted that, by using the series/parallel impedance transformations, the equivalent electrical circuit in Figure 5.5(b) of the two-compartment respiratory system model (cf. Figure 5.5(a)) can be simplified to a form similar to that in Figure D.1(b), which is equivalent to the single-compartment model (cf. Figure D.1(a)). Therefore,



(a) Single-compartment model of respiratory system

(b) Equivalent circuit

Figure D.1 — Single-compartment model of respiratory system and the equivalent electrical circuit with R and C

in what follows, the simulation with one-compartment lung model can be considered without loss of generality.

Let R, C respectively be the equivalent resistance of the airways and the equivalent compliance of lung and chest-walls. The mechanical interaction between the patient's respiratory system and the ventilator is controlled by the following simplified differential equation:

$$P_{\text{aw}} + P_{\text{mus}} = \frac{1}{C}V + R\dot{V} \tag{D.1}$$

where

- P_{aw} is the airway pressure, i.e. the external (ventilator) pressure at the airway.
- P_{mus} is the muscle pressure, resulted by the respiratory muscle. This is internal pressure representing patient's effort.
- V is the lung volume.
- $\dot{V} = \frac{\mathrm{d}V}{\mathrm{d}t}$ is the change rate in volume, i.e. the air flow.

With further ventilatory specifications such as waveform of flow during inspiratory phase, air volume for each inspiration, inspiratory/pause/respiratory time, patient's effort waveform, the whole functioning of the ventilatory support system can be virtualized on computer. Respiratory signals such as flow, volume and pressure can thus be simulated. On the basic of such signals, various detection can be carried out or tested in real-time and in a fully closed-loop.

Simulator with animation and AutoPEEP/Asynchrony detection

As a preliminary work, we implemented such a virtual environment in which the interaction between a ventilator and a patient is simulated. The detection of AutoPEEP/Asynchrony is also integrated in a real-time monitoring situation. The implementation is done in Mathwork's MATLAB. The GUI (Graphical User Interface) is as follows:

The GUI includes a **display** which is used to show the signals, the markers, etc. Any change will be immediately updated on this virtual display for real-time animation. The control panel can be divided into two main sections:

- **Simulation settings** involve both patient's behavior and ventilator's specification. On the one hand, passive mechanical characteristics of patient's respiratory system — including airways resistance (R) and compliance of lungs/chest walls (C) — can be fully specified. This makes it possible to simulate various types of patient. On the other hand, usual ventilatory specifications can be parameterized. These parameters include the breath rate (Fbr) (also called breathing frequency, ventilation rate, respiratory rate), inspiratory-to-expiratory time ratio (I to E), inspiration pause (P) and tidal volume (Vt). Among simulation settings, observation noise level can also be tuned. This option allows us to evaluate the robustness of any processing applied, in particular, the implemented AutoPEEP/asynchrony detectors.

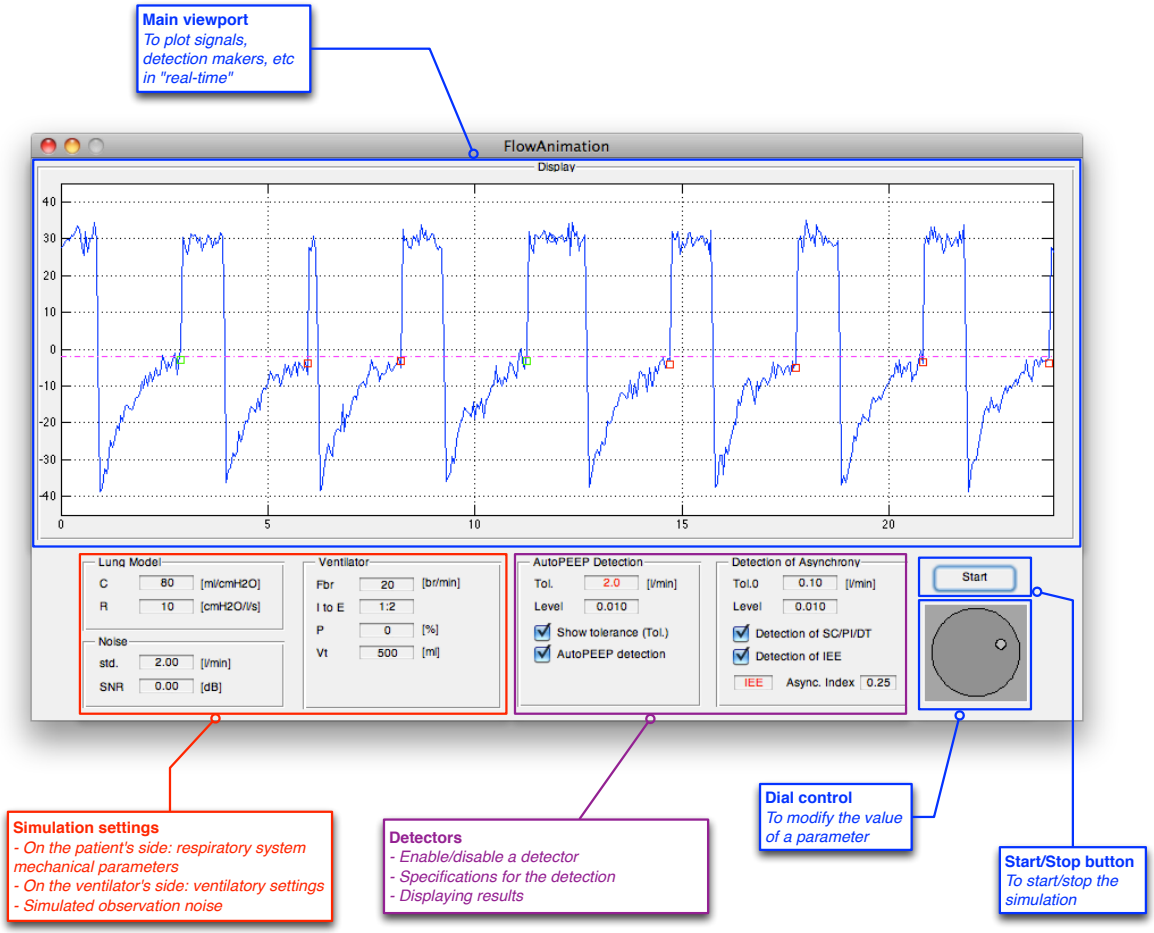


Figure D.2 — Different sections in the GUI (Graphical User Interface) of the implemented virtual ventilatory support simulator, including: the main viewport (display), the simulation settings section, the signal processing blocks (detectors) section and the control buttons.

- **Signal processing blocks** section can integrate any additional functionality. As shown in Figure D.2, the detection of AutoPEEP and asynchrony are included. Tolerance and false-alarm rate level can be specified. For AutoPEEP detection, the results are marked directly on the flow signal in real-time. For the detection of asynchrony (short cycle (SC), prolonged inspiration (PI), double triggering (DT), ineffective effort during expiration (IEE)), the type and the asynchrony index (AI) are provided.

All the aforementioned settings can be modified directly via the corresponding text boxes. Moreover, a **dial control** is also added to facilitate the parameter tuning. For example, in Figure D.2, the tolerance τ for AutoPEEP is being linked to the dial control. Finally, a **start/stop button** is included to start (resp. resume) or stop (resp. pause) the simulation and, also, the animation.

Bibliography

[Alberts et al., 2010] Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2010). *Essential Cell Biology*. Galand Science, third edition.

[Atto and Pastor, 2010] Atto, A. and Pastor, D. (2010). Central limit theorems for wavelet packet decompositions of stationary random processes. *Signal Processing, IEEE Transactions on*, 58(2):896–901.

[Bassevile and Nikiforov, 1993] Bassevile, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes - Theory and Application*. Prentice-Hall, Inc.

[Berman et al., 2000] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1):235–242.

[Berman, 1992] Berman, S. (1992). *Sojourns and extremes of stochastic processes*. Wadsworth and Brooks/Cole.

[Billingsley, 1995] Billingsley, P. (1995). *Probability and Measure*. Wiley, 3rd edition.

[Bishop, 1996] Bishop, C. M. (1996). *Neural Networks for Pattern Recognition*. Oxford University Press, USA, 1 edition.

[Bishop, 2006] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, 2 edition.

[Blanch et al., 2005] Blanch, L., Bernabé, F., and Lucangelo, U. (Jan. 2005). Measurement of air trapping, intrinsic positive end-expiratory pressure, and dynamic hyperinflation in mechanically ventilated patients. *Respiratory Care*, 50:110–123.

[Blanch et al., 2012] Blanch, L., Sales, B., Montanya, J., Lucangelo, U., Garcia-Esquirol, O., Villagra, A., Chacon, E., Estruga, A., Borelli, M., Burgueño, M. J., Oliva, J. C., Fernandez, R., Villar, J., Kacmarek, R., and Murias, G. (Feb 2012). Validation of the better care(®) system to detect ineffective efforts during expiration in mechanically ventilated patients: a pilot study. *Intensive Care Med.*

[Bogan and Thorn, 1998] Bogan, A. and Thorn, K. (1998). Anatomy of hot spots in protein interfaces. *Journal of molecular biology*, 280:1–9.

[Breiman, 1996] Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24:123–140.

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

[Breiman et al., 1984] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC, 1 edition.

[Brinda et al., 2002] Brinda, K. V., Kannan, N., and Vishveshwara, S. (2002). Analysis of homodimeric protein interfaces by graph-spectral methods. *Protein engineering*, 15(4):265–77.

[Cardie, 1993] Cardie, C. (1993). Using decision trees to improve case-based learning. In *Proceedings of the Tenth International Conference on Machine Learning*, pages 25–32. Morgan Kaufmann.

[Chapelle et al., 2006] Chapelle, O., Schölkopf, B., and Zien, A. (2006). *Semi-supervised learning*. MIT Press.

[Chen and Liu, 2005] Chen, X.-W. and Liu, M. (2005). Prediction of protein-protein interactions using random decision forest framework. *Bioinformatics*, 21(24):4394–4400.

[Cho et al., 2009] Cho, K.-i., Kim, D., and Lee, D. (2009). A feature-based approach to modeling protein-protein interaction hot spots. *Nucleic Acids Research*, 37(8):2672–87.

[Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support vector networks. *Machine Learning*, 20:273–297.

[Cosic, 1994] Cosic, I. (1994). Macromolecular bioactivity: is it resonant interaction between macromolecules?-theory and applications. *Biomedical Engineering, IEEE Transactions on*, 41(12):1101–1114.

[Cosic, 1997] Cosic, I. (1997). *The resonant recognition model of macromolecular bioactivity: theory and applications*. Birkhauser Verlag.

[Cosic and Pirogova, 1998] Cosic, I. and Pirogova, E. (1998). Application of ionisation constant of amino acids for protein signal analysis within the resonant recognition model. In *Proceedings of the 20th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, 1998*, volume 2, pages 1072–1075 vol.2.

[Darnell et al., 2007] Darnell, S., Page, D., and Mitchell, J. (2007). Automated Decision-Tree Approach to Predicting Protein-Protein Interaction Hot Spots. *Proteins*, 68(4):813–823.

[Deergha Rao and Swamy, 2008] Deergha Rao, K. and Swamy, M. (2008). Analysis of genomics and proteomics using DSP techniques. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 55(1):370 –378.

[Dempster et al., 1977] Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistic society*, 39, Series B(1):1–38.

[Diaz-Uriarte and Alvarez de Andres, 2006] Diaz-Uriarte, R. and Alvarez de Andres, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 7(1):3.

[Donoho and Johnstone, 1994] Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):pp. 425–455.

[Edward and Cavalli-Sforza, 1965] Edward, A. and Cavalli-Sforza, L. (1965). A method for cluster analysis. *Biometrics*, 21:362–375.

[Fernández-Recio et al., 2004] Fernández-Recio, J., Totrov, M., and Abagyan, R. (2004). Identification of protein-protein interaction sites from docking energy landscapes. *Journal of molecular biology*, 335(3):843–865.

[Fillatre, 2011] Fillatre, L. (2011). *Contributions en détection et classification statistique paramétrique*. HDR - Institut Charles DELAUNAY - Université de technologie de Troyes.

[Fischer et al., 2003] Fischer, T. B., Arunachalam, K. V., Bailey, D., Mangual, V., Bakhru, S., Russo, R., Huang, D., Paczkowski, M., Lalchandani, V., Ramachandra, C., Ellison, B., Galer, S., Shapley, J., Fuentes, E., and Tsai, J. (2003). The binding interface database (BID): a compilation of amino acid hot spots in protein interfaces. *Bioinformatics*, 19(11):1453–1454.

[Freund and Schapire, 1997] Freund, Y. and Schapire, R. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139.

[Gao et al., 2004] Gao, Y., Wang, R., and Lai, L. (2004). Structure-based method for analyzing protein-protein interfaces. *Journal of molecular modeling*, 10(1):44–54.

[Guerois et al., 2002] Guerois, R., Nielsen, J., and Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *Journal of molecular biology*, 320(2):369–387.

[Hampel, 1974] Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):pp. 383–393.

[Harris, 1978] Harris, F. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51 – 83.

[Hartigan, 1975] Hartigan, J. (1975). Clustering algorithms. *Willey*.

[Kawashima et al., 2008] Kawashima, S., Pokarowski, P., Pokarowska, M., Kolinski, A., Katayama, T., and Kanehisa, M. (2008). AAindex: amino acid index database, progress report 2008. *Nucleic acids research*, 36(Database issue):D202–5.

[Keskin et al., 1998] Keskin, O., Bahar, I., Badretdinov, A. Y., Ptitsyn, O. B., and Jernigan, R. L. (1998). Empirical solvent-mediated potentials hold for both intra-molecular and inter-molecular inter-residue interactions. *Protein Science*, 7(12):2578–2586.

[Kortemme and Baker, 2002] Kortemme, T. and Baker, D. (2002). A simple physical model for binding energy hot spots in protein-protein complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22):14116–14121.

[Kortemme et al., 2004] Kortemme, T., Kim, D. E., and Baker, D. (2004). Computational Alanine Scanning of Protein-Protein Interfaces. *Sci. STKE*, 2004(219):pl2–.

[Lai, 1995] Lai, T. L. (1995). Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(4):pp. 613–658.

[Lai and Shan, 1999] Lai, T. L. and Shan, J. (1999). Efficient recursive algorithms for detection of abrupt changes in signals and control systems. *Automatic Control, IEEE Transactions on*, 44(5):952 –966.

[Lee and Richards, 1971] Lee, B. and Richards, F. (1971). The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55(3):379 – 400, IN3–IN4.

[Lefort, 2010] Lefort, R. (2010). *Apprentissage et classification faiblement supervisée : Application en acoustique halieutique*. PhD thesis, TELECOM Bretagne.

[Lehmann and Romano, 2005] Lehmann, E. and Romano, J. (2005). *Testing Statistical Hypotheses*. Springer, 3rd edition edition.

[Leporini and Pesquet, 1999] Leporini, D. and Pesquet, J.-C. (1999). High-order wavelet packets and cumulant field analysis. *Information Theory, IEEE Transactions on*, 45(3):863–877.

[Mahalanobis, 1936] Mahalanobis, P. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2(1):49 – 55.

[Mallat, 1999] Mallat, S. (1999). *A wavelet tour of signal processing*. Academic Press, 2nd edition edition.

[Mulqueeny et al., 2007] Mulqueeny, Q., Ceriana, P., Carlucci, A., Fanfulla, F., Delmastro, M., and Nava, S. (Nov. 2007). Automatic detection of ineffective triggering and double triggering during mechanical ventilation. *Intensive Care Med.*, 33:pp. 2014–2018.

[Neyman and Pearson, 1928] Neyman, J. and Pearson, E. (1928). On the use and interpretation of certain test criteria for purpose of statistical inference. *Biometrika*, 20:175 – 240.

[Nguyen et al., 2012] Nguyen, H.-G., Fablet, R., Ehrhold, A., and Boucher, J. (2012). Keypoint-based analysis of sonar images: Application to seabed recognition. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(4):1171 –1184.

[Ofran and Rost, 2007] Ofran, Y. and Rost, B. (2007). Protein-protein interaction hotspots carved into sequences. *PLoS Comput Biol*, 3(7):e119.

[Opelt et al., 2004] Opelt, A., Fussenegger, M., Pinz, A., and Auer, P. (2004). Weak hypotheses and boosting for generic object detection and recognition. *European Conference on Computer Vision*.

[Pandini et al., 2007] Pandini, A., Mauri, G., Bordogna, A., and Bonati, L. (2007). Detecting similarities among distant homologous proteins by comparison of domain flexibilities. *Protein Engineering Design and Selection*, 20(6):285–299.

[Pastor, 2011] Pastor, D. (2011). Signal norm testing in additive standard gaussian noise. Technical Report RR-2011001-SC, Institut Télécom, Télécom Bretagne.

[Pastor and Atto, 2010] Pastor, D. and Atto, A. (2010). Wavelet shrinkage: from sparsity and robust testing to smooth adaptation. *In Fractals and Related Fields, Eds: J. Barral & S. Seuret*.

[Pastor and Gay, 1995] Pastor, D. and Gay, R. (1995). Décomposition d'un processus stationnaire du second ordre. Propriétés statistiques d'ordre 2 des coefficients d'ondelettes et localisation fréquentielle des paquets d'ondelettes. *Traitement du Signal*, 12(5).

[Pastor and Nguyen, 2012a] Pastor, D. and Nguyen, Q.-T. (2012a). Random distortion testing and optimality of thresholding tests. *IEEE Transactions on Signal Processing*, [submitted].

[Pastor and Nguyen, 2012b] Pastor, D. and Nguyen, Q.-T. (2012b). Testing the mahalanobis distance between a random signal with unknown distribution and a known deterministic model in additive and independent standard gaussian noise: the random distortion testing problem (3rd version). Technical Report RR-2012 04-SC, Institut Télécom, Télécom Bretagne.

[Pastor and Socheleau, 2012] Pastor, D. and Socheleau, F. (2012). Robust estimation of noise standard deviation in presence of signals with unknown distributions and occurrences. *IEEE Transactions on Signal Processing*, 60(4):1545 –1555.

[Quinlan, 1993] Quinlan, R. J. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc.

[Rajamani et al., 2004] Rajamani, D., Thiel, S., Vajda, S., and Camacho, C. J. (2004). Anchor residues in protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31):11287–11292.

[Ramachandran and Antoniou, 2008] Ramachandran, P. and Antoniou, A. (2008). Identification of hot-spot locations in proteins using digital filters. *IEEE Journal on Selected Topics in Signal Processing*, 2(3):378–389.

[Ramachandran et al., 2004] Ramachandran, P., Antoniou, A., and Vaidyanathan, P. (2004). Identification and location of hot spots in proteins using the short-time discrete fourier transform. In *Conference Record of the Thirty-Eighth Asilomar, Conference on Signals, Systems and Computers, 2004*, volume 2, pages 1656–1660.

[Rao, 1948] Rao, C. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, II:50–57.

[Roeseler et al., 2010] Roeseler, J., Michotte, J.-B., and Sottiaux, T. (2010). Patient-ventilator asynchrony during pressure support: Usefulness of the curves displayed by the ventilator. *Réanimation*, 19:pp.62–65.

[Rost, 1999] Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering*, 12(2):85–94.

[Rousseeuw and Croux, 1993] Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):pp. 1273–1283.

[Saeys et al., 2007] Saeys, Y., Inza, I., and Larranaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517.

[Sahu and Panda, 2009] Sahu, S. and Panda, G. (2009). A new approach for identification of hot spots in proteins using s-transform filtering. In *IEEE International Workshop on Genomic Signal Processing and Statistics, 2009 (GENSIPS 2009)*, pages 1–4.

[Seber and Wild, 2003] Seber, G. and Wild, C. (2003). *Nonlinear Regression*. Hoboken, NJ: Wiley-Interscience.

[Serfling, 1980] Serfling, R. (1980). *Approximation theorems of mathematical statistics*. John Wiley & Sons.

[Shivani and Roth, 2002] Shivani, A. and Roth, D. (2002). Learning a sparse representation for object detection. *European Conference on Computer Vision, Springer*.

[Thille et al., 2006] Thille, A., Rodriguez, P., Cabello, B., Lellouche, F., and Brochard, L. (2006). Patient-ventilator asynchrony during assisted mechanical ventilation. *Intensive Care Medicine*, 32:1515–1522. 10.1007/s00134-006-0301-8.

[Thorn and Bogan, 2001] Thorn, K. S. and Bogan, A. A. (2001). ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics*, 17(3):284–285.

[Tuncbag et al., 2009] Tuncbag, N., Gursoy, A., and Keskin, O. (2009). Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics*, 25(12):1513–1520.

[Tuncbag et al., 2010] Tuncbag, N., Keskin, O., and Gursoy, A. (2010). HotPoint: hot spot prediction server for protein interfaces. *Nucleic Acids Research*, 38(suppl 2):W402–W406.

[Ulusoy and Bishop, 2005] Ulusoy, I. and Bishop, C. (2005). Generative versus discriminative methods for object recognition. *Proceedings of the 2005 1EEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:258–265.

[Vaidyanathan and Yoon, 2004] Vaidyanathan, P. P. and Yoon, B.-J. (2004). The role of signal-processing concepts in genomics and proteomics. *Journal of the Franklin Institute*, 341(1-2):111–135.

[Vapnik, 1995] Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.

[Veljkovic and Slavic, 1972] Veljkovic, I. and Slavic, I. (1972). General model of pseudopotentials. *Physical Review Letters*, 29:105–108.

[Vignaux et al., 2009] Vignaux, L., Vargas, F., Roeseler, J., Tassaux, D., Thille, A., Kossowsky, M., Brochard, L., and Jolliet, P. (2009). Patient-ventilator asynchrony during non-invasive ventilation for acute respiratory failure: a multicenter study. *Intensive Care Medicine*, 35:840–846. 10.1007/s00134-009-1416-5.

[Wald, 1943] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482.

[Wells, 1991] Wells, J. (1991). Systematic mutational analyses of protein-protein interfaces. *Methods in enzymology*, 202:390–411.

[Wu et al., 2003] Wu, B., Abbott, T., Fishman, D., McMurray, W., Mor, G., Stone, K., Ward, D., Williams, K., and Zhao, H. (2003). Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. *Bioinformatics*, 19(13):1636–1643.

[Xu and Jelinek, 2007] Xu, P. and Jelinek, F. (2007). Random forests and the data sparseness problem in language modeling. *Computer Speech and Language*, 21(1):105 – 152.

[Younes et al., 2007] Younes, M., Brochard, L., Grasso, S., Kun, J., Mancebo, J., Ranieri, M., J.-C., R., and Younes, H. (Aug 2007). A method for monitoring and improving patient: ventilator interaction. *Intensive Care Med.*, 33:1377–1346.

[Zhang et al., 2008] Zhang, J., Zulkernine, M., and Haque, A. (2008). Random-forests-based network intrusion detection systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(5):649 –659.