Distributed Information Gathering and Estimation in Wireless Sensor Networks

Wenjie Li

NNT : 2016SACLS461

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N° 580
Sciences et technologies de l'information et de la communication

Spécialité de doctorat: Réseaux, Information et Communications

Par

M. Wenjie LI

Collecte et estimation robustes d'information dans un réseau de capteurs sans fils

(Distributed Information Gathering and Estimation in Wireless Sensor Networks)

Thèse présentée et soutenue à Gif sur Yvette, le 15 novembre 2016

Composition du Jury :

Mme LAROCHE Béatrice	Directeur de recherche (INRA)	Présidente
M. FAGNANI Fabio	Professor (Politecnico di Torino)	Examinateur
M. DARDARI Davide	Associate Professor (University of Bologna)	Examinateur
M. RICHARD Cédric	Professeur (Université de Nice Sophia Antipolis)	Rapporteur
M. GORCE Jean-Marie	Professeur (INSA Lyon)	Rapporteur
M. KIEFFER Michel	Professeur (Université Paris-Sud)	Directeur de thèse
Mme BASSI Francesca	Enseignant-Chercheur (ESME-Sudria)	Co-encadrant de thèse

Acknowledgements

Firstly, I would like to give my thanks to my advisors, Prof. Michel Kieffer and Prof. Francesca Bassi for their solid support. Without their guidance this dissertation could not have been finished. It has been an honor for me to work with them and I really expect further collaboration in future.

Secondly, I would like to thank all the colleagues that I had a great opportunity to work with in the project NEWCOM#. Thanks to Prof. Davide Dardari, Prof. Gianni Pasolini and Prof. Laura Galluccio for their inspiring discussions and collaboration. I have passed six wonderful and unforgettable months in Cesena and Bologna, thanks Dr. Vincenzo Zambianchi and Alex Calisti for their warm reception. I would like to thank all the colleagues and secretaries of both LSS and University of Bologna, for their friendly assistant.

Then, I would like to thank my reviewers for the time they have dedicated for the reading of this PhD thesis and the development of their reports.

Finally, I would like to thanks all my friends and family for their strong love and encouragement. I thank my mother Chunping Huang for everything, that cannot be expressed by words. I thank my wife Xi Gong, she makes my life really happy and colorful.

Contents

Résumé en Français

1. Contexte

Les réseaux de capteurs sans fils (RCSFs) [Ver+10] suscitent un intérêt croissant depuis une vingtaine d'années. Ces réseaux permettent de prélever, traiter, et diffuser les mesures de phénomènes physiques, tels que la température, la pression, ou l'hygrométrie sur de vastes zones géographiques. Leur domaine d'application s'étend de la surveillance d'environnements extérieurs au contrôle d'ateliers industriels.

Cette thèse aborde quelques problèmes en lien avec la collecte et l'acheminement efficace de données ainsi qu'avec l'estimation distribuée dans un RCSF.

1.1. Codage réseau pour la collecte distribuée de données corrélées

Afin de faciliter le déploiement et d'améliorer la robustesse des RCSFs, la communication entre les nœuds peut être effectuée en l'absence de point d'accès désigné et de structure hiérarchique. Au niveau de la couche réseau, la diffusion des mesures vers tous les nœuds peut être réalisée, par exemple à l'aide d'un protocole d'inondation (*flooding*), ou en utilisant un protocole exploitant le *codage réseau* (CR) [Ahl+00], afin de réduire le trafic. Lorsqu'un schéma de CR est utilisé, les données mesurées sont quantifiées et placées par chaque capteur au sein de paquets qu'il envoie vers ses voisins. Ces derniers réalisent des combinaisons linéaires des paquets reçus et des paquets qui contiennent leurs données propres. Ces combinaisons linéaires sont ensuite réémises. Le plus souvent, les coefficients des combinaisons linéaires sont choisis aléatoirement, et sont envoyés dans chaque en-tête de paquet. Un nœud du réseau souhaitant effectuer un traitement doit prélever un certain nombre de paquets contenant des combinaisons linéaires indépendantes des paquets contenant les données mesurées par les capteurs.

Sans exploiter la corrélation entre données, il faut recevoir autant de paquets contenant des combinaisons linéaires indépendantes qu'il y a de capteurs dans le réseau. Dans de nombreuses situations pratiques, les mesures obtenues par le RCSF sont spatialement et temporellement corrélés. Cette corrélation peut être exploitée pour effectuer un décodage, comme cela se fait dans [IKAA12; Bas+12; RAG12; BTF12]. Le premier objectif de cette thèse est d'étudier le nombre de combinaisons linéairement indépendantes qui sont nécessaires et suffisantes pour récupérer parfaitement les données que les capteurs mesurent.

1.2. Auto-évaluation distribuée

Les nœuds d'un RCSF peuvent être équipés de capteurs ayant des caractéristiques différentes en terme de bruit de mesure ou de capacité de traitement. Chaque nœud n'a pas nécessairement une

connaissance complète des caractéristiques du bruit de ses capteurs, par exemple, biaisé et non biaisé, voir [Chi+11]. Il peut être intéressant d'estimer ces caractéristiques. Plus largement, dans des applications de type *crowdsensing* [Guo+15], où les réseaux de capteurs sont constitués de nœuds participant de manière opportuniste à la collecte de données, la fourniture de services d'acquisition fiables exige des mesures de bonne qualité. De telles mesures ne sont pas toujours disponibles dans tous les appareils participant à la collecte, sachant que la précision d'acquisition dépend fortement des capteurs embarqués. Il est donc important d'identifier les appareils fournissant les mesures de la meilleure qualité. Ce problème est considéré, par exemple, en utilisant des mécanismes basés sur la réputation [KM14a; KM14b; Ren+15]. Dans ces systèmes les données sont collectées à partir de smartphones, sur la base d'un mécanisme d'enchères. Les tâches d'acquisition sont attribuées par une autorité centrale aux agents selon leur niveau de réputation, afin de maximiser l'utilité pour le réseau. Ces mécanismes basés sur la réputation sont efficaces pour l'évaluation des agents, mais ont besoin d'une centralisation des données et d'une autorité de supervision.

Nous supposerons dans cette thèse que chaque agent ou nœud du réseau ne connait pas les caractéristiques de son dispositif d'acquisition et qu'il souhaite auto-évaluer ces caractéristiques. La seconde partie de cette thèse sera dédiée à la conception et à l'analyse d'algorithmes d'*auto-évaluation distribués* (AED), qui permettent à chaque nœud d'un réseau d'auto-évaluer les caractéristiques de ses capteurs. Plus généralement, ces outils permettent à des agents d'un réseau d'évaluer leur niveau d'expertise dans la réalisation de certaines tâches (mesure, décision, classification...).

Nous commencerons par un problème simple d'AED, à savoir la Détection Distribuée des nœuds possédant des capteurs Défaillants (DDD) d'un réseau. C'est un problème d'AED où seuls deux niveaux de performance des capteurs sont considérés: bon ou défectueux. Les capteurs défectueux produisent fréquemment des mesures aberrantes, c'est-à-dire, des mesures anormales qui ne peuvent pas être justifiées par les simples effets du bruit d'acquisition [BL94]. La présence de nœuds équipés de capteurs défectueux peut considérablement dégrader les performances d'un RCSF dédié à une tâche d'estimation distribuée. Il est donc très important de détecter ces nœuds efficacement (avec des coûts de communication et de calcul faibles), idéalement de manière distribuée.

1.3. Prise en compte d'une connectivité intermittente

Les réseaux tolérants aux déconnections (RTDs) sont des réseaux dont la topologie est dynamique et le degré de connectivité très faible. Les communications entre nœuds du réseau sont alors très limitées et de courte durée. Cette intermittence rend les problèmes d'estimation et d'auto-évaluation de performance beaucoup plus complexes que dans des réseaux plus classiques [HCY11; SRF14; Zhu+14; GLG16; Pan+15].

Ce type de contrainte de communication apparaît dans des RCSFs où afin d'économiser de l'énergie, les nœuds sont la plupart du temps en veille et ne communiquent que de manière sporadique. Le modèle de RTD est alors très adapté dans ce contexte [Fal03; KAF12].

La dernière partie de cette thèse est consacrée au développement et à l'analyse d'algorithme de DDD pour des RTDs. Ces résultats sont ensuite généralisés à l'AED en considérant un nombre fini de niveaux d'expertise. Les agents du réseau seront supposés avoir des interactions sporadiques et

seulement de pair à pair, comme c'est généralement le cas dans des RTDs, ou dans des réseaux où les échanges sont effectués via *gossiping* [FZ08; Dim+10].

2. Codage réseau et compression distribuée dans un RCSF

Dans un RCSF, le codage conjoint source-réseau de données produites par des sources corrélées peut être effectué en utilisant des techniques empruntées au *compressed sensing* (CS) [Don06]. L'objectif du CS est l'estimation d'un vecteur $\boldsymbol{\theta} \in \mathbb{R}^N$ à partir de mesures consistant en des combinaisons linéaires des composantes de $\boldsymbol{\theta}$ dont le nombre M peut être nettement inférieure à N. Si $\boldsymbol{\theta}$ est *k-sparse* dans une base connue, la reconstruction de $\boldsymbol{\theta}$ peut alors être effectuée presque sûrement à partir $\mathcal{O}(k \log(N))$ combinaisons linéaires en utilisant des algorithmes de type *basis poursuit* [CRT06; CT06a]. Des résultats similaires peuvent être obtenus lorsque $\boldsymbol{\theta}$ est seulement compressible [Don06], avec une qualité de reconstruction correspondant à celle permise par l'observation directe des k coefficients de plus grandes valeurs absolue de $\boldsymbol{\theta}$ dans un domaine transformé.

La caractéristique principale du CS est que les combinaisons linéaires ne doivent pas nécessairement être adaptées au signal à acquérir. Il suffit que des conditions, dites de *restricted isometry* sur la matrice de mesure soient satisfaites [CT06a; CT05]. Par ailleurs, le CS est robuste à la présence de bruit dans les mesures [CT05; HN06].

L'estimation de $\boldsymbol{\theta}$ a été abordée dans un contexte bayésien [JXC08]. Le vecteur à estimer est alors vu comme une réalisation d'un vecteur aléatoire $\boldsymbol{\Theta}$, dont une distribution *a priori* traduit le caractère *sparse* ou la compressibilité des composantes. Dans les implémentations pratiques, l'estimation de $\boldsymbol{\theta}$ à partir des combinaisons linéaires de ses composantes peut être réalisée en exploitant des modèles graphiques statistiques [Mon12], et des algorithmes de type *belief propagation* (BP) [KFL01], voir [BSB10] pour des matrices de mesure déterministes et [BM11] pour des matrices de mesure aléatoires.

Lorsque l'on cherche à combiner CS et CR dans des RCSF, on est rapidement confronté au fait que les mesures des capteurs sont à valeurs réelles ou sont quantifiées et que les opérations de CR le sont en général dans des corps finis. En conséquence, la matrice de CR, qui joue le rôle de la matrice de mesure dans le CS classique, a des éléments dans un corps fini. Il est alors difficile d'exploiter les algorithmes de reconstruction efficace du CS. Une première solution à ce problème est proposée par [FM11], où le CR est réalisé dans le corps des réels. Ce cadre est également adopté dans [LCJ15], où la reconstruction des données est effectuée progressivement pour réduire le délai de décodage. Dans [NL14], la matrice des combinaisons linéaires résultant du CR dans le corps des réels est quantifiée avant transmission. Le prix à payer par toutes ces techniques est de plus grands en-têtes et une incompatibilité avec le CR classique.

L'analyse de la performance des estimateurs de $\boldsymbol{\theta}$ a été abordée dans [DM09; RAG12], et [SL13]. Le travail de [SL13] n'a pas considéré d'*a priori* Bayésien, et suppose un niveau connu de *sparsité* de $\boldsymbol{\theta}$. Un décodage idéal via une minimisation de norme ℓ_0 est supposé. Les conditions nécessaires et suffisantes pour l'estimation exacte de $\boldsymbol{\theta}$ sont dérivés, en fonction de la taille de $\boldsymbol{\theta}$, de son niveau de *sparsité*, du nombre de mesures, et de la *sparsité* de la matrice d'acquisition. Les résultats numériques montrent que les conditions nécessaires et suffisantes coïncident, lorsque la taille de $\boldsymbol{\theta}$

augmente asymptotiquement. Un contexte Bayésien est considéré dans [DM09] et [RAG12]. Dans [DM09], une distribution *a priori* induit la *sparsité* de la réalisation de $\mathbf{\Theta}$, dont les éléments sont supposées statistiquement indépendants. En utilisant la méthode des types [Csi98], l'exposant d'erreur de reconstruction en utilisant une minimisation ℓ_0 est obtenu en l'absence de bruit de mesure, et l'exposant d'erreur de reconstruction en utilisant un décodeur minimisant l'entropie empirique est détaillé pour des mesures bruitées. Dans [RAG12], certains modèles de corrélation spécifiques (entre paire de composantes, entre groupe de composantes) de $\mathbf{\Theta}$ sont considérés. Les exposants d'erreur pour un décodage MAP sont proposés, à nouveau en cas l'absence de bruit de mesures.

Le chapitre 2 de cette thèse est consacré à l'étude de l'efficacité de compression de données corrélées provenant d'un RCSF et acheminées vers un point de collecte à l'aide du codage réseau linéaire aléatoire (CRLA). Nous supposons un contexte Bayésien et nous considérons un décodage au sens du maximum *a posteriori*. Le modèle de système considéré est représenté dans la figure 1. On suppose que les données prélevées appartiennent à $\mathbb{F}_Q^N$, le corps fini à Q éléments. Le vecteur aléatoire $\mathbf{\Theta}^N \in \mathbb{F}_Q^N$ représente le vecteur de source et $\mathbf{x}^N \in \mathbb{F}_Q^N$ est une copie de $\mathbf{\Theta}^N$ corrompu par le bruit de mesure. Le chapitre 2 considère un modèle de source plus général que celui considéré dans la littérature, à savoir un modèle stationnaire et ergodique, de taux d'entropie $\mathcal{H}(\Theta)$. Chaque capteur du réseau prélève une composante de $\mathbf{x}^N$, la quantifie à l'aide d'un quantificateur à Q intervalles de quantification. Chaque nœud réalise un CRLA du paquet contenant cette valeur quantifiée et des paquets reçu des autres nœuds.

Les effets du CRLA peuvent être modélisés par la multiplication de $\mathbf{x}^N$ par une matrice aléatoire $\mathbf{A} \in \mathbb{F}_Q^{M\times N}$. Les éléments de $\mathbf{A}_i$, la i-ème colonne de $\mathbf{A}$, sont supposés être indépendants et identiquement distribués (iid) dans $\mathbb{F}_Q$ selon

$$p_{\mathsf{A}_i}(0) = 1-\gamma_i, \qquad p_{\mathsf{A}_i}(q) = \frac{\gamma_i}{Q-1} \text{ pour tout } q \in \mathbb{F}_Q \setminus \{0\} \tag{1}$$

où γ_i est le facteur de sparsité de $\mathbf{A}_i$. Les valeurs de γ_i pour les différentes colonnes ne sont pas nécessairement identiques. Nous supposons qu'un bruit additif $\mathbf{u}^M \in \mathbb{F}_Q^M$ affecte les paquets reçus. Le vecteur d'entrée du décodeur, notée $\mathbf{y}^M \in \mathbb{F}_Q^M$, peut être représenté comme

$$\mathbf{y}^M = \mathbf{A}\mathbf{x}^N + \mathbf{u}^M. \tag{2}$$

Le décodeur observe la réalisation $\mathbf{y}^M$ de $\mathbf{y}^M$ et connait parfaitement la réalisation $\mathbf{A}$ de $\mathbf{A}$, par exemple, grâce aux en-têtes des paquets, voir [CWJ03] et [Jaf+09]. L'estimation au sens du maximum a posteriori $\hat{\boldsymbol{\theta}}^N$ de la réalisation de $\mathbf{\Theta}^N$ est alors

$$\hat{\boldsymbol{\theta}}^N = \arg\max_{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N \mid \mathbf{y}^M, \mathbf{A}\right). \tag{3}$$

En nous inspirant du travail de [SL13], nous cherchons à obtenir des conditions nécessaires et suffisantes pour la reconstruction presque sûrement exacte de $\boldsymbol{\theta}$, lorsque sa taille augmente asymptotiquement. En utilisant des outils tels que le théorème de Shannon-McMillan-Breiman, ou la

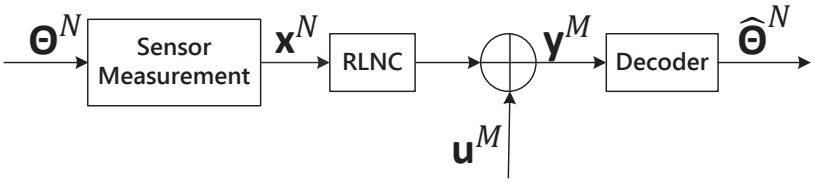

FIGURE 1: Schéma du modèle de système

typicité faible, nous avons montré que le taux de compression optimal dépend du niveau de corrélation de source. On considère $\mathcal{H}(\Theta, \mathsf{x})$ comme le taux d'entropie jointe entre la valeur de la source Θ et sa mesure x et $H(p_{\mathsf{u}})$ comme l'entropie du bruit additif. Asymptotiquement, il faut que

$$\frac{M}{N} > \frac{\mathcal{H}(\Theta, \mathsf{x})}{\log Q - H(p_{\mathsf{u}})}, \tag{4}$$

et que l'entropie conditionnelle $H\left(\mathbf{\Theta}^N \mid \mathbf{x}^N\right)$ tende vers zéro pour que l'erreur d'estimation puisse devenir arbitrairement faible. De (4), on peut déduire qu'il n'est possible de reconstruire parfaitement $\boldsymbol{\theta}^N$ en présence de bruit de mesure que lorsque $H(p_{\mathsf{u}}) < \log Q$. De plus, on a aussi trouvé une condition suffisante sur les facteurs de *sparsity* γ_i

$$\min_{i=1\ldots N} \gamma_i > 1 - 2^{-H(p_{\mathsf{u}})}, \tag{5}$$

pour assurer la reconstruction parfaite. Cette condition indique que la matrice de codage peut être arbitrairement *sparse*, dans le cas sans bruit.

Les hypothèses relatives aux caractéristiques de la source et de la matrice de codage dans notre travail sont beaucoup plus réalistes que celles qui ont été considérées préalablement. De plus, aucune analyse n'a été effectuée précédemment tenant compte d'un bruit d'acquisition. Nos résultats pour le cas sans bruit sont compatibles avec ceux présentés dans [SL13]; en outre, nous pouvons formellement prouver la convergence asymptotique des conditions nécessaires et suffisantes, et étendre les bornes sur le facteur de sparsité de la matrice de codage en présence de bruit additif. Les analyses asymptotiques avec décodage MAP, pour le cas sans bruit et en présence de bruit additif, sont compatibles avec les résultats obtenues dans [DM09], respectivement sous décodage par minimisation de norme ℓ_0 et par le décodage minimisant l'entropie empirique. Les exposants d'erreur pour le décodage MAP des sources corrélées dans le cas sans bruit sont compatibles avec ceux présentés dans [RAG12], et sont ici étendus au cas d'une source ergodique et en présence des deux types de bruits.

Les résultats en lien avec ce sujet ont fait l'objet de:

- W. Li, F. Bassi, and M. Kieffer, "Sparse Random Linear Network Coding for Data Compression in WSNs", *Proc. IEEE International Symposium on Information Theory, ISIT 2016*, Barcelona, Spain, 2016. [LBK16]

- W. Li, F. Bassi, and M. Kieffer, "Robust Bayesian compressed sensing over finite fields: asymptotic performance analysis", in preparation. [LBK14]

3. Détection distribuée des nœuds défaillants dans un RCSF

De nombreux travaux ont été entrepris en lien avec la problématique de la détection des nœuds équipés de capteurs défectueux au sein d'un RCSF, voir par exemple [ZMH10; MK13; Don+14]. L'article [MK13] propose une classification des différentes techniques de détection. Certaines techniques reposent uniquement sur les données provenant du voisinage direct de chaque nœud [CKS06], d'autres utilisent également les données provenant des nœuds plus éloignés [DP+15; Bra+13]. Certaines techniques reposent sur un modèle paramétrique du phénomène observé par les capteurs pour identifier des mesures aberrantes [Ise05; Lo+13]. Les techniques, qui ne sont pas à base de modèles, exploitent principalement des propriétés statistiques des mesures, telles que la corrélation spatiale et temporelle. A titre d'exemple, l'algorithme de [CKS06] permet à chaque nœud d'estimer son état en comparant ses mesures locales avec celles de ses voisins. Dans une première phase seule une tendance (bon, défectueux, probablement bon, ou probablement défectueux) est déterminée. Dans une deuxième phase, chaque nœud recueille les tendances de ses voisins et les associée à ses mesures pour obtenir une évaluation plus fiable. Dans [Din+05] une comparaison entre la mesure locale et la valeur médiane des mesures des voisins est réalisée. Dans [PK15], une version modifiée d'un test reposant sur un écart de plus de trois écart-types du bruit est considéré pour identifier les nœuds produisant des mesures de très grande variance. Des algorithmes itératifs sont proposés dans [GXL07; Ji+10], où respectivement les critères de la médiane pondérée et de la moyenne pondérée sont considérés. Dans les deux cas, le test local pondère les mesures des voisins par leur niveau de confiance qui sont obtenu à partir d'une détection préalable, dans l'hypothèse d'une défaillance permanent des nœuds. Dans [LC08], l'algorithme utilise la redondance temporelle afin de lutter contre les défauts transitoires de détection et de communication. L'algorithme adaptatif proposé dans [Cho+09] ajuste le seuil de décision à chaque étape pour améliorer sa précision de détection. Dans [Bra+13], une fonction générique d'identification de mesures aberrantes est supposée être disponible à chaque nœud. Les résultats fournis par cette fonction à chaque nœud, ainsi que les mesures sont échangées pour permettre à l'ensemble du réseau d'identifier toutes les *mesures aberrantes* dans un ensemble de données. Dans [DP+15], une approche bayésienne distribuée est proposée pour détecter les mesures aberrantes dans un grand ensemble de données collectées par un RCSF.

En général, la disponibilité de plus de données à un nœud de référence facilite la détection et l'identification de mesures aberrantes, au prix d'un coût de communication et d'un délai de détection qui augmente avec la quantité de données. Une approche intéressante pour réduire le nombre de tests est l'approche par *group testing* (GT), qui est une méthode statistique permettant d'identifier un petit nombre de mesures aberrantes dans un grand ensemble de données en effectuant un nombre limité de tests élémentaires, appelés test locaux de détection de mesures aberrantes (TLD), impliquant des sous-ensembles de données [DH93]. Ces tests élémentaires sont seulement en mesure de déterminer si les mesures aberrantes sont présents dans l'ensemble considéré. Un algorithme distribué de GT [Che+11], qui étend les résultats obtenus en utilisant le GT centralisé, est proposée dans [TTF13], dans l'hypothèse d'un taux de capteurs défectueux assez faible. La solution proposée dans [TTF13] présente cependant un inconvénient important dans les en-têtes de communication nécessaires, pour l'indexation des nœuds et pour la diffusion des résultats des tests

à l'ensemble du RCSF.

Malgré l'abondante littérature sur le sujet de la détection de capteurs défaillants, il y a encore de nombreux problèmes ouverts identifiés dans [MK13; Don+14]. Les nœuds d'un RCSF sont généralement limités en termes de mémoire, puissance de calcul et énergie disponible. La plupart des systèmes de DDD traditionnels ont accordé peu d'attention à ces problèmes. Ils sont généralement gourmands en énergie et en mémoire, et ont souvent besoin d'un nombre important de communications.

Par ailleurs, les travaux précédents se sont souvent concentrés sur l'amélioration des TLD. Ceux-ci sont ensuite analysés en considérant des études de cas spécifiques par des simulations. La relation entre les paramètres des algorithmes de détection, comme le seuil de décision de défaillance d'un capteur et les performances de ces algorithmes est difficile à prévoir. A notre connaissance, peu de travaux ont cherché à mettre les techniques de DDD dans un cadre générique et à fournir une analyse de l'équilibre ou de la stabilité des algorithmes itératifs de DDD.

Au Chapitre 3, nous considérons un ensemble de nœuds $\mathcal{S}$ déployés uniformément et de façon indépendant dans un plan. Un nœud ne peut que communiquer avec les nœuds à portée de communication. On définit $\mathcal{D} \subseteq \mathcal{S}$ comme l'ensemble des nœuds équipés de capteurs défectueux et $\mathcal{G} = \mathcal{S} \setminus \mathcal{D}$ comme l'ensemble des nœuds avec des bons capteurs. Le statut du nœud $i \in \mathcal{S}$ est représenté par $\theta_i \in \{0, 1\}$. On a $\theta_i = 1$ si $i \in \mathcal{D}$ et $\theta_i = 0$ si $i \in \mathcal{G}$.

Nous considérons des TLD capable uniquement de détecter la présence de mesures aberrantes dans un ensemble de mesures $\mathbf{M}_{\mathcal{A}}$ fournis par un ensemble de nœuds $\mathcal{A}$, sans nécessairement être capable d'identifier lesquelles des données sont erronées. Cette hypothèse élargit la gamme des TLDs applicables aux TLD n'ayant besoin que de peu de mesures (typiquement 2 ou 3). Les TLDs sont caractérisés par $q_{\mathrm{D}}(\mathbf{M}_{\mathcal{A}})$ et $q_{\mathrm{FA}}(\mathbf{M}_{\mathcal{A}})$, les probabilités de détection de mesures aberrantes, dans le cas où $\mathcal{A} \cap \mathcal{D} \neq \emptyset$ et de fausse alarme, lorsque $\mathcal{A} \cap \mathcal{D} = \emptyset$, respectivement. Par ailleurs, pour simplifier de l'analyse, nous considérons uniquement les TLDs vérifiant les propriétés suivantes

- *P1*). $q_{\mathrm{D}}(\mathbf{M}_{\mathcal{A}}) = q_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}})$ et $q_{\mathrm{FA}}(\mathbf{M}_{\mathcal{A}}) = q_{\mathrm{FA}}(n_{\mathrm{g}})$, où $n_{\mathrm{g}} = |\mathcal{A} \cap \mathcal{G}|$ et $n_{\mathrm{d}} = |\mathcal{A} \cap \mathcal{D}|$;
- *P2*). $q_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}})$ et $q_{\mathrm{FA}}(n_{\mathrm{g}})$ augmentent avec n_{g} et n_{d}.

Il est très facile de construire de tels TLDs, par exemple dans un contexte d'estimation à erreurs bornées [Mil+96; Jau+01].

Le chapitre 3 de cette thèse, propose deux algorithmes de DDD inspirés des techniques de GT: une version avec une seule décision (SDDD) et une autre itérative (IDDD). La seconde version est mieux adaptée aux RCSF possédant une grande proportion de capteurs défectueux. Dans les deux cas, on découple le TLD, qui fournit seulement une indication (vrai, faux) de la présence de mesures aberrantes dans un ensemble de mesures, de la décision finale sur le statut de nœud (bon, défectueux).

Les algorithmes proposés consistent de deux étapes : une détection locale de la présence de mesures aberrantes en utilisant un TLD, et la décision. Dans la première étape, chaque nœud i recueille des données provenant de son voisinage, effectue le TLD pour déterminer si des mesures aberrantes sont présentes parmi les données recueillies. On introduit $\mathcal{N}_i$ comme l'ensemble de voisins de nœud i et $\mathcal{U}_i = \mathcal{N}_i \cup \{i\}$. Idéalement (sans aucune perte de paquets et d'erreurs de transmission),

le TLD effectué par le nœud i, dont le résultat est y_i, implique les données fournies par les nœuds $\mathcal{U}_i$. Chaque nœud diffuse le résultat de son TLD à ses voisins et calcule la somme des y_j qu'il a reçus, c'est-à-dire $\sum_{j\in\mathcal{U}_i} y_j$. La nature en deux étapes de l'algorithme rend la phase de décision indépendante de la nature du TLD considéré. Nous formulons les hypothèses que le TLD devrait satisfaire pour permettre de prouver l'existence d'un état d'équilibre. Par ailleurs, il est possible de caractériser analytiquement les performances de l'algorithme et ainsi d'ajuster la valeur de ses paramètres. Cet aspect est l'une des problématiques ouvertes mentionnées dans [ZMH10; MK13].

Dans la version avec une seule décision de l'algorithme DDD proposé, les nœuds réalisent L itérations de la première étape et ensuite estiment le statut de ses capteurs dans une phase de décision unique sur la base des résultats des TLDs effectuées et reçus des nœuds voisins

$$\widehat{\theta}_i = \begin{cases} 1 \ \text{(défectueux)} & \text{si } \sum_{\ell=1}^{L} \sum_{j\in\mathcal{U}_i} y_j^{(\ell)} \geqslant \gamma L \left|\mathcal{U}_i\right|, \\ 0 \ \text{(bon)} & \text{sinon.} \end{cases} \tag{6}$$

Cette version ne fonctionne bien que dans le cas où le rapport des nœuds dont les capteurs sont défectueux est très faible. Sinon, la probabilité qu'un bon nœud soit considéré comme défectueux, notée par P_{FA}, peut être élevée.

Dans la version itérative de l'algorithme DDD, chaque nœud est autorisé à prendre des décisions temporaires $\widehat{\theta}_i^{(\ell)}$ sur le statut de son capteur. Les nœuds avec $\widehat{\theta}^{(\ell-1)} = 1$ restent silencieux au cours de l'itération ℓ. Toute décision peut être mise à jour au cours des prochaines itérations. Les nœuds réalisent L_1 itération de ces traitements. Ensuite, les nœuds avec $\widehat{\theta}^{(L_1)} = 0$ mettent en œuvre l'algorithme DDD avec décision unique et L_2 itérations. Les nœuds pour lesquels $\widehat{\theta}^{(L_1)} = 1$ sont déterminés comme défectueux et ne participent pas aux dernières L_2itérations. Quelques exemples numériques sont représentées sur Figure 2, dont le détail de la simulation peut être trouvé dans la thèse au Chapitre 3.5. Les L_1 premières itérations aident à diminuer P_{FA} et les L_2 itérations suivante améliorent P_{D}, qui est la probabilité qu'un nœud défectueux soit détecté.

Il est intéressant d'étudier la dynamique de la proportion de nœuds dans chaque état $(\theta, \widehat{\theta})$, avec $\theta, \widehat{\theta} \in \{0, 1\}$. Des conditions suffisantes à satisfaire par le TLD ont été mises en évidence pour assurer l'existence d'un équilibre des algorithmes DDD.

Un avantage des algorithmes DDD proposés est que, lorsque les conditions de communication sont parfaites, les nœuds ne doivent que transmettre le résultat binaire de son TLD à chaque itération. Cependant, comme chaque nœud diffuse des paquets au cours de l'algorithme, les problèmes de canal et les pertes de paquets ne peuvent être négligés. Les algorithmes DDD initiaux peuvent encore fonctionner, mais avec des performances dégradées. Une alternative consiste, en plus de la transmission de y_i, pour chaque nœud i, de transmettre également un ensemble $\mathcal{V}_i \subseteq \mathcal{U}_i$ afin de signaler les indices des voisins qui ont été impliqués dans son TLD. Chaque nœud i peut décider de prendre en compte le y_j reçu en vérifiant si $i \in \mathcal{V}_j$.Une discussion détaillée sur le trafic généré par l'algorithme DDD est présenté dans la thèse pour des situations différentes. Il est intéressant de voir s'il est nécessaire de transmettre $\mathcal{V}_i$. La figure 3 compare des algorithmes itératifs avec ou sans la transmission de $\mathcal{V}_i$, ainsi que l'algorithme de référence [LC08], avec différentes proportions de nœuds défectueux. Dans cet exemple, la transmission de $\mathcal{V}_i$ n'est pas très importante tant que

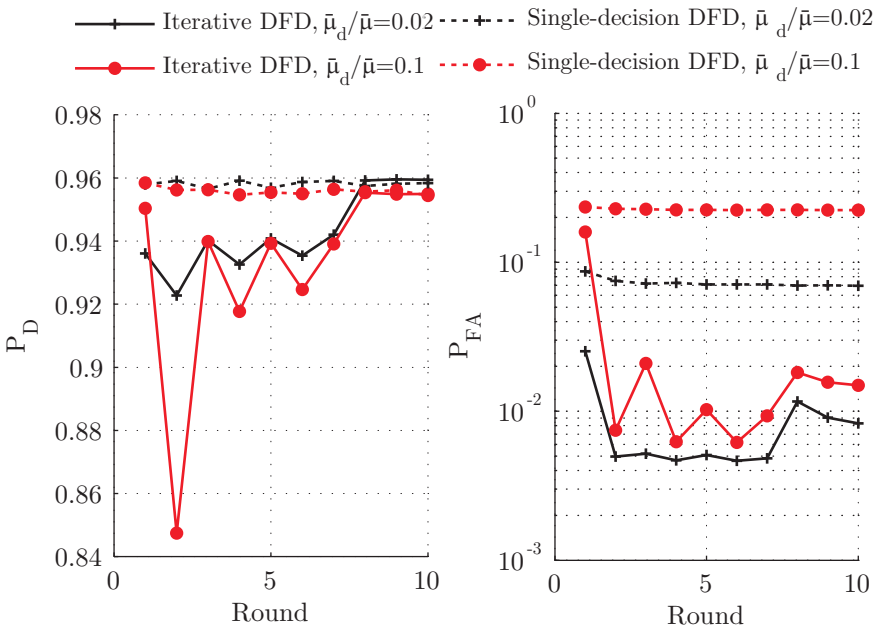

FIGURE 2: P_{D} (à gauche) et P_{FA} (à droite) en fonction du nombre d'itérations, avec un taux des nœuds défectueux $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$, $L = 10$ pour le DDD avec décision unique; $L_1 = 7$ et $L_2 = 3$ pour le DDD itératif.

$P_{\text{success}} \geqslant 0.6$, où P_{success} est la probabilité qu'un nœud réussisse à accéder au canal et à transmettre son paquet.

Pour étudier l'impact des protocoles de communication et des conditions de propagation réelles, l'algorithme proposé a également été mis en œuvre sur la plate-forme *Data Sensing and Processing Testbed* (*DataSens*), qui est une partie des installations expérimentales disponibles dans EuWIn@CNIT/Bologne. La plate-forme adoptée est composé de 20 ou de 41 nœuds sans fil de type EMB-2530PA [Ins]. La pile logicielle TIMAC, conforme au standard IEEE 802.15.4, a été utilisée pour l'implantation des algorithmes. Les résultats obtenus sont ensuite comparés avec ceux de la simulation pour évaluer l'efficacité de l'algorithme. L'algorithme proposé réalise de bonne performance (le rapport des nœuds ayant une mauvaise estimation du statut est inférieur à 5%) dans un court délai (1 seconde), même dans un environnement de communication instable.

En résumé, les principales contributions du Chapitre 3 sont

- La proposition d'algorithmes de DDD de complexité réduite, reposant sur des TLDs très génériques qui ont le potentiel de fonctionner avec très peu de mesures ;
- L'obtention des conditions à satisfaire par un TLD et par la densité des nœuds du réseau pour faire en sorte que l'algorithme DDD itératif atteigne un équilibre ;
- Une caractérisation analytique des performances dans des conditions de canal réalistes, liées aux pertes de paquets et les problèmes d'accès au canal à la couche MAC, ce dernier problème étant rarement pris en compte dans d'autres travaux ;
- La mise en œuvre de l'algorithme proposé dans un réseau de capteurs réel pour vérifier l'efficacité des algorithmes.

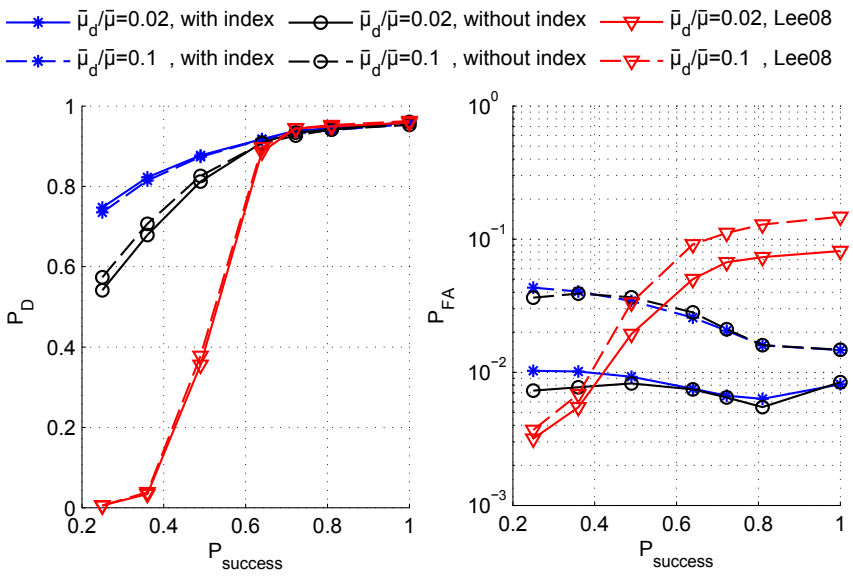

FIGURE 3: Comparaison des algorithmes itératifs avec ou sans la transmission de $\mathcal{V}_i$, ainsi que l'algorithme de référence [LC08], avec $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$.

Les résultats en lien avec ce sujet sont dans les publications:

- W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests", *IEEE transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp 29-48, 2016. [Li+16a]
- W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Low-Complexity Distributed Fault Detection for Wireless Sensor Networks", *Proc. IEEE International Conference on Communications, ICC 2015*, London, UK, 2015, pp. 6712–6718. [Li+15b]
- W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Iterative Distributed Outlier Detection for Wireless Sensor Networks: Equilibrium and Convergence Analysis", *Proc. IEEE Conference on Decision and Control, CDC 2015*, Osaka, Japan, 2015, pp. 3050-3056. [Li+15a]
- W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Impact of Channel Access Issues and Packet Losses on Distributed Outlier Detection within Wireless Sensor Networks", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2016*, Shanghai, China, 2016. [Li+16c]

4. Détection distribuée des nœuds défaillants dans un RTD

Le problème de DDD est beaucoup moins étudié dans le contexte des RTDs que dans les RCSFs génériques. Lorsque un peu de mesures sont disponibles, les TLDs classiques ne sont pas très efficaces: par exemple, il est très difficile, à partir de deux mesures seulement fournies par deux nœuds différents, de déterminer lequel de ces nœuds est défectueux. Ceci est une situation typique dans RTDs lorsque deux nœuds se rencontrent, prennent des mesures, et partagent ces mesures.

Une application directe des algorithmes DDD classiques dans le contexte des RTDs peut donc être tout à fait inefficace.

Un problème lié a été considéré par [Zhu+12], dans le contexte des RTDs véhiculaires (VRTDs). Un grand nombre de capteurs sont fixes et liés à l'infrastructure routière. Certains véhicules, appelés *mobile carriers* (MC), recueillent des données issues de ces capteurs. Les capteurs ne peuvent communiquer qu'avec les MCs de leur voisinage. Un MC a besoin de recueillir suffisamment de mesures pour effectuer un test permettant de déterminer les mesures produites par des capteurs défectueux. Une fois qu'un nœud est jugé défectueux par un MC, il est ajouté à la liste noire. Les MCs fournissent des informations aux capteurs sur leur statut et échangent aussi leurs listes noires afin d'accélérer la détection des capteurs défaillants.

Dans [Pen+14], le problème de détection de nœuds infectés par des logiciels malveillants et formant un RTD est considéré. Chaque nœud détermine après la rencontre avec un autre nœud si celui-ci a effectué des actions suspectes (essai de transmission de logiciel malveillant). Les nœuds détectés comme infectés sont isolés.

La gestion de confiance ou de la réputation des nœuds est un autre aspect important pour aider un RTD à résister à des attaques. Ainsi, [AF12] fournit un mécanisme itératif de gestion de la confiance pour lutter contre les attaques byzantines où des nœuds sont totalement contrôlés par un adversaire. Un modèle de confiance pour les réseaux de capteurs acoustiques sous-marins est présenté dans [Han+15] pour prendre en compte plusieurs paramètres de confiance tels que la confiance de lien, la confiance des données, et la confiance du nœud. La gestion des niveaux de confiance est également très importante dans des application de type Sensing as a Service (SaaS), voir par exemple [She+13].

Le chapitre 4 présente un algorithme entièrement distribué et facilement implémentable pour permettre à chaque nœud d'un RTD de déterminer si ses propres capteurs sont défectueux. Les conditions de communication des RTDs imposent une structure d'algorithme DDD complètement différente de celle introduite au Chapitre 3, même si les deux algorithmes utilisent des TLD.

Le statut d'un nœud est représenté par la variable $\theta \in \{0,1\}$ où $\theta = 0$ désigne un nœud normal et $\theta = 1$ indique un nœud défectueux. La proportion de nœuds avec $\theta = 0$ et $\theta = 1$ sont p_0 et p_1, avec $p_0 + p_1 = 1$. Comme dans [Zhu+10; GLG16; HO+15; Pan+15], nous supposons que l'intervalle du temps entre deux réunions successives suit une distribution exponentielle de paramètre λ. De plus, nous supposons que chaque réunion implique seulement deux nœuds. Lorsque plus de deux nœuds se rencontrent au même temps, le traitement est effectué paire par paire. Lorsque 2 nœuds ont une réunion, ils peuvent échanger leurs données et effectuer un TLD identique à celui considéré au chapitre 3. Dans le contexte des RTDs, un TLD implique $n_0 \in \{0,1,2\}$ nœuds normaux et $n_1 \in \{0,1,2\}$ nœuds défaillants, avec $n_0 + n_1 = 2$. On suppose en outre que $q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(1,1) \leqslant q_{\mathrm{D}}(0,2)$, qui est raisonnable: la probabilité de détection de la présence d'un nœud défaillant augmente avec le nombre de nœuds défaillants participant au TLD.

L'idée principale de l'algorithme RTD-DDD proposé est simple: chaque nœud gère deux compteurs $c_{\mathrm{m},i}(t)$ et $c_{\mathrm{d},i}(t)$ initialisés à 0 à l'instant $t = 0$. En utilisant $c_{\mathrm{m},i}(t)$, le nœud i compte le nombre de TLDs qu'il a effectués. Avec $c_{\mathrm{d},i}(t)$, il compte le nombre de TLD ayant détecté la présence de mesures aberrantes. Chaque nœud met à jour l'estimée de son statut en utilisant la règle de décision

$$\widehat{\theta}_i(t) = \begin{cases} 1 \text{ (défectueux)}, & \text{si } c_{\mathrm{d},i}(t)/c_{\mathrm{m},i}(t) \geqslant \nu, \\ 0 \text{ (bon)}, & \text{sinon.} \end{cases} \tag{7}$$

où ν est un seuil de décision constant. Lorsqu'un nœud i avec $\widehat{\theta}_i(t) = 1$ rencontre un nœud j, le nœud i n'envoie pas ses mesures locales vers nœud j, mais il effectue encore un TLD et met à jour ses compteurs locaux si $\widehat{\theta}_j(t) = 0$.

Le chapitre 4 étudie le comportement de l'algorithme RTD-DDD à l'aide des modèles de Markov et d'outils empruntés à la théorie du contrôle et à l'étude de la dynamique des populations. Pour ce faire, la croyance de chaque nœud sur l'état de ses capteurs est quantifiée. On introduit $\mathbf{x}_i(t) = (\theta_i, c_{\mathrm{m},i}(t), c_{\mathrm{d},i}(t))$ comme l'état du nœud i à l'instant t. Parmi les nœuds dont le statut est θ, on définit

- $X_\theta^{\ell,k}(t)$: la proportion des nœuds dans l'état (θ, ℓ, k), avec $c_{\mathrm{m},i}(t) = \ell$ et $c_{\mathrm{d},i}(t) = k$;
- $p^{\theta\widehat{\theta}}(t)$: la proportion des nœuds avec le statut réel θ et le statut estimé $\widehat{\theta}$.

On remarque que p^{11} représente le taux de détection, *i.e.*, la proportion de mauvais capteurs qui ont effectivement identifié leurs capteurs comme défectueux, et p^{01} représente le taux de fausse alarme, *i.e.*, la proportion de nœuds qui croient que leurs capteurs normaux sont en fait défectueux.

Pour limiter le nombre d'états possibles, on ne considère que les M dernières réunions au cours desquelles le nœud i a effectué un TLD. Ainsi, $c_{\mathrm{d},i}(t) \leqslant c_{\mathrm{m},i}(t) \leqslant M$. L'évolution de $\mathbf{x}_i(t)$ suit deux chaînes de Markov indépendantes, suivant la valeur de θ_i. La dynamique des proportions $X_\theta^{\ell,k}(t)$ est décrite à l'aide d'un ensemble d'équations différentielles. On a montré qu'il existe toujours un certain équilibre du système dynamique pour $\nu \in [0,1]$. Au point d'équilibre, on a $\overline{X}_\theta^{\ell,k} = 0$, $\forall k \leqslant \ell < M$ et $\overline{X}_\theta^{M,k}$ s'exprime de la manière suivante

$$\overline{X}_\theta^{M,k} = \binom{M}{k} (h_\theta)^k (1 - h_\theta)^{M-k},$$

avec

$$h_0 = \frac{p_0 q_{\mathrm{FA}}(2)\,\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\,\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}, \quad h_1 = \frac{p_0 q_{\mathrm{D}}(1,1)\,\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\,\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}},$$

dans lesquelles $\overline{p}^{00}$ et $\overline{p}^{10}$ peuvent être obtenus en résolvant

$$\begin{cases} \overline{p}^{00} = \sum\limits_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\mathrm{FA}}(2)\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}} \right)^k \left(\frac{p_0(1-q_{\mathrm{FA}}(2))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(1,1))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}} \right)^{M-k}, \\ \overline{p}^{10} = \sum\limits_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\mathrm{D}}(1,1)\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}} \right)^k \left(\frac{p_0(1-q_{\mathrm{D}}(1,1))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(0,2))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}} \right)^{M-k}. \end{cases} \tag{8}$$

Des conditions suffisantes sur les paramètres de décision pour assurer l'unicité d'un équilibre de l'algorithme RTD-DDD sont ensuite fournies. Lorsque M tend vers l'infini, on peut prouver que tous les $\overline{p}^{01}$ et $\overline{p}^{10}$ tendent vers 0, ce qui signifie que chaque nœud est capable d'estimer correctement son statut, si la valeur de M est suffisamment grande.

Les valeurs exactes de $\overline{X}_\theta^{M,k}$s sont difficile à obtenir. Cependant, on peut obtenir les approximations suivantes

$$\widetilde{X}_0^{M,k} = \binom{M}{k} \left(q_{\mathrm{FA}}(2)\right)^k \left(1 - q_{\mathrm{FA}}(2)\right)^{M-k}, \quad \widetilde{X}_1^{M,k} = \binom{M}{k} \left(q_{\mathrm{D}}(1,1)\right)^k \left(1 - q_{\mathrm{D}}(1,1)\right)^{M-k}, \quad (9)$$

puis l'approximation de $\overline{p}^{\theta 0}$ est

$$\widetilde{p}^{\theta 0} = \sum_{k:k/M<\nu} \widetilde{X}_\theta^{M,k}.$$

Les valeurs réelles des $\overline{p}^{00}$ and $\overline{p}^{10}$ sont très proches de $\widetilde{p}^{11}$ and $\widetilde{p}^{01}$, dans la région où $\overline{p}^{11}$ est proche de 1. Ces approximations fournissent des lignes directrices pour choisir correctement la valeur de ν.

En simulation, on considère tout d'abord certains modèles idéaux de communication pour vérifier la validité des résultats théoriques obtenus. L'algorithme RTD-DDD est ensuite simulé en considérant des RTD dont les rencontres entre nœuds ont été stockées dans des bases de données expérimentales bien étudiés, voir le projet Haggle [Sco+09] et le projet du MIT Reality Mining Projet [EP06]. Plus précisément, nous avons extrait de ces bases les instants de rencontres entre nœuds. Dans chaque test, N_D nœuds sont tirés aléatoirement comme défaillants et l'algorithme RTD-DDD est lancé pour permettre à tous les nœuds de déterminer leur statut. Des simulations de Monte-Carlo sont effectuées 500 fois pour les bases de données *Reality* et *Infocom05*. En haut de la figure 4, chaque point suer une ligne indique un nœud actif (qui est en contact avec un autre nœud). Ceci montre l'évolution de la fréquence des contacts au cours du temps. L'évolution de p^{10} et p^{01} est en bas de la figure 4. Tous les p^{10} et p^{01}, obtenus par les deux bases de données, diminuent jusqu'à 10^{-3} après une durée suffisamment longue. On observe également que la vitesse de convergence de p^{10} et p^{01} est fortement liée à la fréquence des contacts. La figure 5 représente les états à l'équilibre $\overline{X}_\theta^{M,k}$s obtenus en utilisant les bases de données *Reality* et *Infocom05*, et ainsi que ceux résultant de l'approximation (9). Il y a une excellente coïncidence entre les valeurs à l'équilibre prédite par la théorie et celles obtenues par simulation.

Enfin, on caractérise l'influence des paramètres sur les performances de l'algorithme RTD-DDD. Pour montrer l'efficacité de l'algorithme proposé, on considère $q_{\mathrm{D}}(0,2) = 0.9$ et $M = 10$ fixés. Pour $p_1 = 0.1$ et $p_1 = 0.5$, on s'intéresse aux $\overline{p}^{10}$ et $\overline{p}^{01}$ réalisables pour $0 \leqslant q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(0,2)$ et $q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(1,1) \leqslant q_{\mathrm{D}}(0,2)$. Quatre domaines sont considérés: Dans le domaine 3: tous les $\overline{p}^{10}$ et $\overline{p}^{01}$ sont inférieures à 10^{-3}; dans le domaine 2: tous les $\overline{p}^{10}$ et $\overline{p}^{01}$ sont inférieurs à 10^{-2}; dans le domaine 1: tous les $\overline{p}^{10}$ et $\overline{p}^{01}$ sont inférieurs 10^{-1}; dans le domaine 0: $\overline{p}^{10}$ ou $\overline{p}^{01}$ peuvent être plus grands que 10^{-1}. La figure 6 montre la partition du triangle $(q_{\mathrm{D}}(1,1), q_{\mathrm{FA}}(2))$ en quatre domaines, représentés par des couleurs différentes. Le ratio de nœuds défectueux dans le réseau n'a pas d'impact significatif sur la performance à l'équilibre, même quand 50% des nœuds sont défectueux.

Les résultats en lien avec ce sujet sont dans les publications:

- W. Li, L. Galluccio, M. Kieffer, and F. Bassi, "Distributed Faulty Node Detection in DTNs", *Proc. International Conference on Computer Communication and Networks, ICCCN 2016*, Hawaii, USA, 2016. [Li+16b]

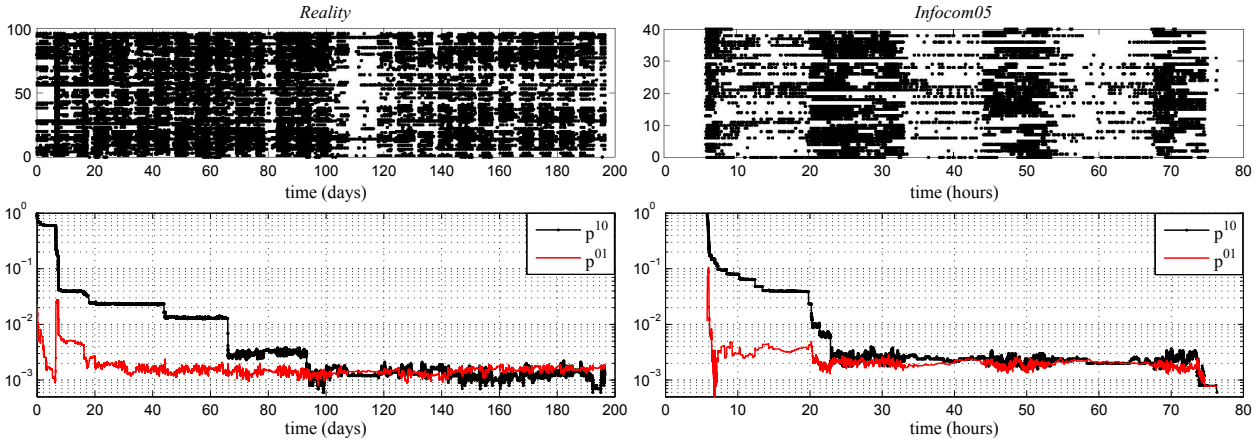

FIGURE 4: Indices des nœuds actifs à différents moments (en haut) et l'évolution de p^{01} et p^{10} (en bas) obtenus en utilisant les bases de données *Reality* avec $N_D = 20$ (à gauche) et *Infocom05* avec $N_D = 10$ (à droit), avec $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$, et $\nu = 0.4$.

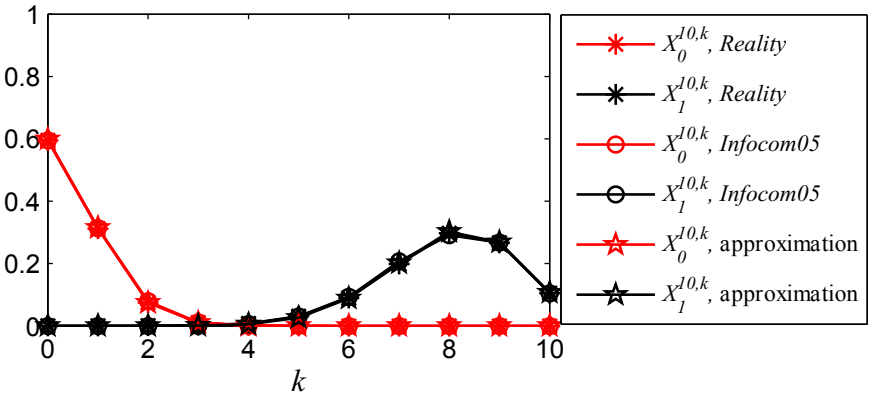

FIGURE 5: Comparaison de $\overline{X}_{\theta}^{10,k}$ à l'équilibre obtenu en utilisant les bases de données *Reality* et *Infocom05*, et prédite par l'approximation (9).

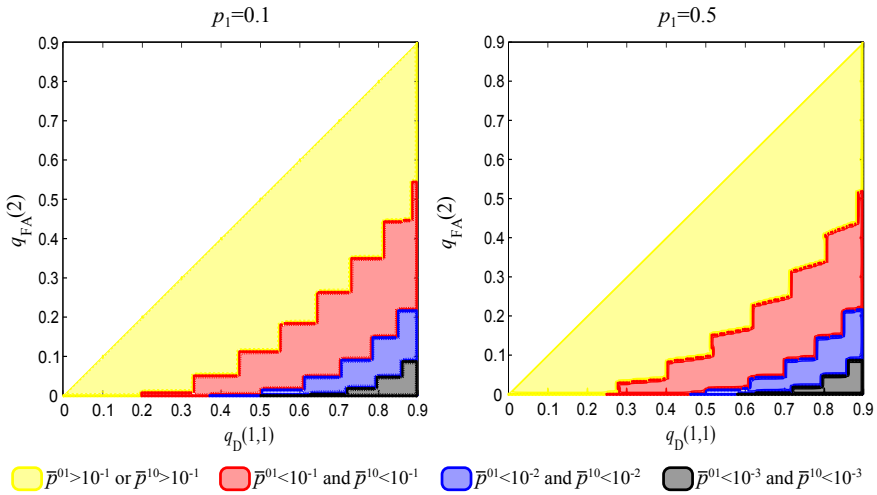

FIGURE 6: Réalisable $\overline{p}^{10}$ et $\overline{p}^{01}$ pour différentes valeurs de la paire $(q_{\mathrm{D}}(1,1), q_{\mathrm{FA}}(2))$ lorsque $p_1 = 0.1$ (à gauche) et $p_1 = 0.5$ (à droit).

- W. Li, L. Galluccio, M. Kieffer, and F. Bassi, "Distributed Faulty Node Detection in Delay Tolerant Networks: Design and Analysis", *IEEE transactions on Mobile Computing*, submitted. [Li+16d]

5. Auto-évaluation distribuée

Le chapitre 5 considère un groupe d'agents collaborant afin d'exécuter une tâche donnée (acquisition, détection, classification, voir par exemple [Luo+07; Ang+09; SBW16]), pour laquelle chaque agent a potentiellement un niveau d'expertise différent. Ce niveau d'expertise peut être liée, par exemple à la qualité des capteurs ou à la puissance de calcul que chaque agent a à sa disposition. Initialement, nous supposons que chaque agent ne connait pas son niveau d'expertise, ou s'il le connait, ne sais pas se situer vis-à-vis des autres agents. L'objectif de l'auto-évaluation distribuée (AED) est de permettre à chaque agent, par l'intermédiaire d'échanges de pair-à-pair avec d'autre agents, d'évaluer son niveau d'expertise par rapport à celui d'autre agents.

Ce type de problème est important des dans applications de type sensing as a service (SaaS). Permettre à chaque agent d'être conscient de son niveau d'expertise permet de fournir des mesures ou des décision avec des niveaux de confiance associés. Des mécanismes de réputations sont mis en œuvre dans ces dispositifs. En cas de défaillance temporaire, cela permet à un agent de ne pas transmettre de données, plutôt que de corrompre sa réputation. Cela permet également aux meilleurs agents de mieux valoriser la qualité de leurs mesures ou de leurs décisions, sans avoir besoin de l'aide d'un organe de traitement central.

Des problèmes simples d'AED ont été abordés dans le passé par [Chi+11; FFR14], où chaque nœud d'un RCSF doit estimer la valeur d'un paramètre constant à partir de mesures scalaires bruitées, ainsi que le biais [Chi+11] ou le niveau de la variance [FFR14] de son bruit de mesure, tous les deux supposés ne prendre que deux valeurs distinctes. Les nœuds sont partitionnés en deux groupes dans [Chi+11], chaque groupe étant caractérisé par l'absence ou la présence de biais de mesure. L'algorithme proposé nécessite un consensus [HM09] avec des échanges de type pair-à-pair par *gossiping*, ainsi que le classement distribué des nœuds [FZ08] selon la valeur de leur mesure. Chaque nœud est alors capable de décider si son bruit de mesure est biaisé. Les partitions considérées dans [FFR14] dépendent des valeurs possibles de la variance du bruit de mesure.

Dans le chapitre 5 un nombre fini de niveaux d'expertise est considéré. Ce niveau d'expertise, représenté par $\theta_i \in \Theta = \{1, \dots, K\}$ pour l'agent i est supposé constant et décrit sa capacité à effectuer une certaine tâche. Nous supposons que les niveaux d'expertise vont en décroissant avec θ. Ainsi, $\theta = 1$ indique le niveau d'expertise maximal et $\theta = K$ le niveau minimal. L'algorithme d'AED implique des interactions entre agents de type pair-à-pair et des tests de comparaison (TC). Un TC implique deux agents et est capable d'estimer l'agent ayant le niveau d'expertise le plus élevé, mais n'est pas capable de quantifier la différence de niveau d'expertise. Ainsi, le TC renvoie $y_{i,j} = 1$ lorsqu'il est vraisemblable que l'agent i a un niveau d'expertise supérieur à celui de l'agent j. Il renvoie $y_{i,j} = 0$ sinon. Tout comme les TLDs, les TC sont caractérisés par leurs propriétés statistiques $q(\theta_i, \theta_j) = \mathbb{P}\{Y_{i,j} = 1 | \theta_i, \theta_j\}$. Seules ces propriétés sont importantes dans la suite. Le TC peut donc être très générique.

Tout comme pour l'algorithme de RTD-DDD, chaque agent i utilise deux compteurs, $c_{\mathrm{t},i}(t)$ et $c_{\mathrm{b},i}(t)$ pour compter le nombre de TCs qu'il a effectués et le nombre de TCs ayant produit $y_{i,\cdot} = 1$. Intuitivement, un agent dont le niveau d'expertise est θ a plus de chances d'avoir un plus grand rapport $c_{\mathrm{b},i}(t)/c_{\mathrm{t},i}(t)$ qu'un agent avec un niveau d'expertise θ' lorsque $\theta < \theta'$. On peut alors introduire une partition de l'intervalle $[0,1]$ en K intervalles de décision $[\nu_k, \nu_{k-1})$ avec $\nu_0 = 1$ et $\nu_K = 0$, chacun correspondant à un niveau d'expertise estimé

$$\widehat{\theta}_i\,(t) = k \text{ si } c_{\mathrm{b},i}(t)/c_{\mathrm{t},i}(t) \in [\nu_k, \nu_{k-1})\,,\ k = 1\ldots, K. \tag{10}$$

Nous considérons que lorsque deux agents i et j se rencontrent, ils commencent à échanger leur niveau d'expertise estimé. L'agent i demande à poursuivre l'interaction avec une probabilité $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ et l'agent j avec une probabilité $\alpha\left(\widehat{\theta}_j, \widehat{\theta}_i\right)$. L'interaction se poursuit lorsqu'au moins l'un des agents la demandent. Seuls les agents qui demandent à poursuivre l'interaction effectuent le TC et exploitent son résultat. Les valeurs de $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ sont des paramètres d'ajustement pour optimiser les performances de l'algorithme d'AED.

L'efficacité de l'algorithme d'AED proposé est mesurée par le *taux de décision correcte* (TDC) et par le *taux de décision erronée* (TDE), c'est-à-dire, la proportion des agents qui auto-évaluent correctement ou non leur niveau d'expertise. L'analyse de l'algorithme d'AED est réalisée en supposant une population bien mélangée d'agents pour lesquels l'intervalle de temps entre deux rencontres est distribué exponentiellement. Ce modèle permet de décrire l'évolution des proportions d'agents ayant des estimées similaires dans leur expertise par des équations d'état à temps continu.

Comme au chapitre 4, on introduit $X_\theta^{\tau,\beta}(t)$ comme la proportion des agents avec $c_{\mathrm{b},i}(t) = \beta$ et $c_{\mathrm{t},i}(t) = \tau$, parmi les agents dont le niveau d'expertise est θ. Ici, p_θ représente la proportion d'agents dont le niveau d'expertise est θ ; $p^{\theta\widehat{\theta}}$ représente la proportion d'agents de niveau d'expertise réel θ et de niveau estimé $\widehat{\theta}$. Afin de simplifier l'analyse, chaque nœud ne considère que les M dernières résultats des TCs.

L'analyse du comportement de l'algorithme d'AED pour des valeurs arbitraires de α est difficile. On se focalise sur deux cas particuliers ici.

- Dans le *cas I*, une rencontre entre deux agents conduit toujours à une interaction, $\alpha(k_1, k_2) = 1$, $\forall k_1, k_2$. Dans ce cas, on peut obtenir l'expression exacte des proportions d'agents à l'équilibre, $\overline{X}_\theta^{\tau,\beta} = 0$, $\forall \tau < M$ et

$$\overline{X}_\theta^{M,\beta} = \binom{M}{\beta}\left(\sum_{k\in\Theta} p_k q\,(\theta, k)\right)^{\beta}\left(1 - \sum_{k\in\Theta} p_k q\,(\theta, k)\right)^{M-\beta}. \tag{11}$$

- *Dans le cas II*, une rencontre entre deux agents se poursuit par une interaction uniquement lorsqu'un agent i rencontre un agent j estimant que son niveau d'expertise est le meilleur,

$$\alpha\,(k_1, k_2) = \begin{cases} 1 & \text{si } k_2 = 1, \\ 0 & \text{sinon.} \end{cases} \tag{12}$$

Dans ce cas, l'expression explicite de $\overline{X}_{\theta}^{M,\beta}$ n'est pas facile à écrire. On a trouvé cependant une condition suffisante pour assurer l'existence d'un équilibre $\overline{X}_{\theta}^{M,\beta}$. Ceci permet ensuite d'obtenir une approximation

$$\widetilde{X}_{\theta}^{M,\beta} = \binom{M}{\beta} \left(q\left(\theta,1\right)\right)^{\beta} \left(1-q\left(\theta,1\right)\right)^{M-\beta}, \tag{13}$$

qui peut être vue comme une extension du résultat (9).

En utilisant (11) ou (13), on peut optimiser les seuils de décision introduits dans (10). Par exemple, pour maximiser la somme de TDC , *i.e.*, $\sum_{\theta\in\Theta}\overline{p}^{\theta\theta}$ ou $\sum_{\theta\in\Theta}\widetilde{p}^{\theta\theta}$, les valeurs de ν_{θ} doivent satisfaire

$$\nu_{\theta} = \frac{\log\left(\frac{1-a_{\theta}}{1-a_{\theta+1}}\right)}{\log\left(\frac{a_{\theta+1}}{a_{\theta}}\frac{1-a_{\theta}}{1-a_{\theta+1}}\right)}, \qquad \forall\theta = 1\ldots\left(K-1\right), \tag{14}$$

où $a_{\theta} = \sum_{k\in\Theta} p_k q\left(\theta,k\right)$ pour le cas I et $a_{\theta} = q\left(\theta,1\right)$ pour le cas II.

Pour la partie expérimentale, on utilise à nouveau les deux bases de données expérimentales, *Reality* et *Infocom05*. Le modèle de TC est décrite en détail au chapitre 5.4, page 103-104. La figure (7) montre l'évolution de $P_e^{\theta} = 1 - p^{\theta\theta}$ pour le cas II. Pour les deux bases de données, P_e^{θ} décroit aux alentours de 10^{-2} pour toutes les valeurs de $\theta \in \Theta$ après un temps suffisant. La figure 8 représente les valeurs de $\overline{X}_{\theta}^{M,\beta}$ obtenues en utilisant les bases de données *Reality* et *Infocom05*, et aussi par l'approximation (13). Dans ce cas, il y a à nouveau une excellente coïncidence entre la théorie et la simulation. Cependant, les résultats pour le Cas I, également fourni dans la thèse, ne sont pas aussi bons. Ceci illustre l'importance du choix de α. Un travail important reste à faire pour analyser le comportement de l'algorithme d'AED avec des probabilités d'interaction quelconques.

Une partie des résultats en lien avec ce sujet sont dans les publications:

- W. Li, F. Bassi, L. Galluccio, and M. Kieffer, "Self-Rating in a Community of Peers", *IEEE Conference on Decision and Control, CDC 2016*, Las Vegas, USA, 2016. [Li+16e]
- W. Li, F. Bassi, L. Galluccio, and M. Kieffer, "Peer-assisted Individual Assessment in a Multi-agent system, *Automatica*, submitted. [Li+16f]

6. Perspectives

Plusieurs problèmes ouverts peuvent être considérés à la suite de cette thèse.

6.1. Codage réseau linéaire aléatoire

Au Chapitre 2, nous avons analysé les limites en termes de compression de données dans un RCSF en présence de bruit, en considérant une matrice de codage avec différents niveaux de sparsité.

Pour assurer le taux de compression optimal, nous avons proposé une condition suffisante sur le niveau de sparsité minimal de la matrice de codage réseau. Un travail reste à faire sur la recherche des conditions que doit satisfaire le niveau de sparsité moyen. La réciproque de la preuve que nous

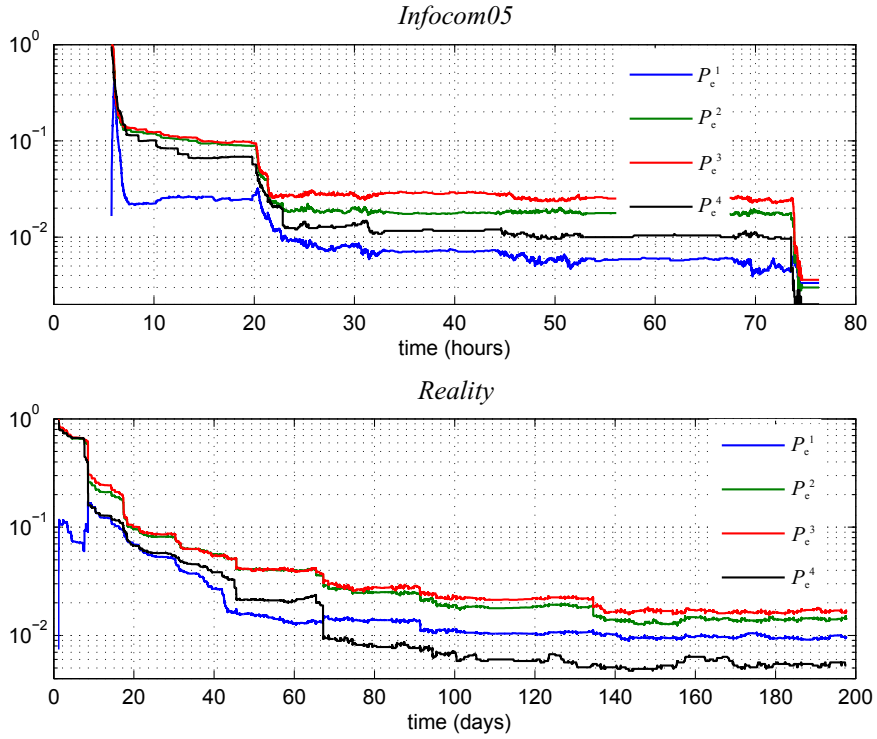

FIGURE 7: Évolution de $P_e^\theta = 1 - p^{\theta\theta}$ obtenues en utilisant les bases de données *Reality* (en haut) et *Infocom05* (en bas).

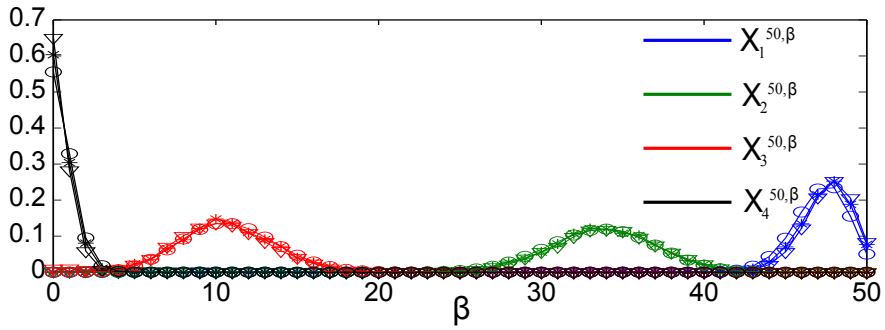

FIGURE 8: Valeurs de $\overline{X}_\theta^{50,\beta}$ à la fin de la simulation: les cercles sont les valeurs théoriques obtenues à partir de ((13)), les triangles sont les résultats de la simulation à l'aide du *Reality*, et les croix sont des résultats de simulation à l'aide du *Infocom05*.

avons proposée repose principalement sur l'inégalité de Fano. Aucune condition nécessaire sur la sparsité de la matrice de codage n'a été proposée. Il est important de chercher ce type de résultat, afin de déterminer si les conditions nécessaires et suffisantes sur la sparsité peuvent coïncider.

Compte-tenu des résultats obtenus sur le facteur de sparsité, un problème intéressant est de voir comment configurer un RCSF pour permettre de satisfaire les conditions sur la sparsité. Ainsi, lorsque les nœuds sont uniformément distribués sur un plan avec une densité donnée, il s'agit de déterminer comment les nœuds doivent réaliser les combinaisons linéaires des paquets.

6.2. Détection distribuée de défauts dans un RTD

Dans ce chapitre, nous avons supposé que le statut des nœuds reste constant sur l'intervalle de temps que dure la phase de détection de défauts. Il serait intéressant d'introduire un statut des nœuds variant au cours du temps, ce qui est plus réaliste. Ainsi, un nœud normal peut devenir défectueux lorsque sa puissance est faible, un nœud défaillant peut redevenir normal lorsque son capteur défectueux est remplacé, ou lorsqu'il recharge sa batterie. Dans cette situation, des probabilités de transition entre statuts doivent être introduites et les chaines Markov conditionnées par les deux statuts (voir, par exemple, la figure 4.1) ne seront plus indépendantes. L'analyse est plus complexe dans ce cas.

Les techniques développées pour l'analyse de la dynamique de l'algorithme de détection de défauts peut être utile afin analyser d'autres types de problèmes d'auto-estimation. Par exemple, cet outil peut être appliqué à la détection d'infection des nœuds par des logiciels malveillants. Les nœuds du réseau doivent à nouveau être rationnels et souhaiter connaître leur statut (infecté par un virus ou sains). Lors d'une rencontre, nous supposons que chaque nœud dispose d'un test permettant d'évaluer si le nœud rencontré a tenté de lui transmettre un logiciel malveillant. La détection de logiciels malveillants est plus difficile que la détection du capteur défectueux, à cause de la nature des logiciels malveillants: un nœud normal peut être infecté par un nœud infecté lors de leur interaction. Il est donc indispensable de considérer l'évolution temporelle du statut des nœuds. Le modèle de propagation des logiciels malveillants (voir [KSA12] par exemple) peut être considéré et introduit dans notre modèle de Markov pour analyser la dynamique. Il y aura un compromis entre la proportion des nœuds infectés actifs et la proportion des nœuds normaux qui sont mal détectés comme infectés.

6.3. Auto-évaluation distribuée

Les résultats présentés au chapitre 5 en lien avec l'auto-évaluation distribuée du niveau d'expertise des agents d'une communauté sont obtenus en considérant deux cas particuliers de probabilités d'interaction entre agents: soit les agents interagissent à chaque rencontre, soit ils n'interagissent que s'ils rencontrent un agent estimant avec le niveau d'expertise maximal. Un travail important reste à faire pour analyser le comportement de l'algorithme d'auto-évaluation distribué avec des probabilités d'interaction quelconques. L'existence et l'unicité de l'équilibre de la dynamique doit être analysé dans le cas général. L'objectif est de déterminer les probabilités d'interaction optimales afin de maximiser la précision de l'estimation du niveau d'expertise pour différents scénarios applicatifs.

On peut également chercher, sous la contrainte d'un taux de mauvaise décision donné, de minimiser le nombre moyen d'interactions entre agents. Ceci peut être intéressant lorsque les interactions ont un coût en termes de temps ou d'énergie.

Ce problème est lié aux problèmes de bandits manchots (BM) [AB10], dont l'objectif est d'estimer la statistique des différentes machines afin de choisir la machine maximise la gain. Dans ce scénario, il n'y a pas de contrainte lors de la phase d'apprentissage sur le choix de la machine à tester. Dans le cas de l'auto-évaluation distribuée, chaque agent doit décider, lors d'une rencontre avec un autre agent s'il continue à interagir avec cet agent ou s'il préfère attendre de rencontrer un agent qui lui permettra de gagner une meilleure connaissance de son niveau d'expertise. Cette contrainte supplémentaire, liée aux RTD, rend l'analyse plus complexe.

Un autre problème similaire est le classement des agents en utilisant une comparaison de type pair-à-pair [Hec+16]. Lorsque le nombre de groupes de niveaux d'expertise similaire est égal au nombre d'agents, le problème d'auto-évaluation distribué devient un problème de classement distribué. Ce problème est délicat en l'absence d'organe de traitement central permettant d'ordonnancer les rencontres entre agents. En l'absence d'organe de traitement central, le seul choix pour accélérer l'algorithme est d'agir sur les probabilités d'interaction, et éventuellement d'autoriser ces probabilités à évoluer au cours du temps. Ainsi, en début de classement, une interaction pourra avoir lieu à chaque rencontre entre agents. Au bout d'un certain temps, la probabilité d'interaction peut évoluer comme une fonction décroissante avec la différence entre les niveaux d'expertise estimés des deux agents. Il s'agit de trouver la probabilité d'interaction assurant le meilleur compromis entre le nombre d'interactions nécessaires et la vitesse de convergence de l'algorithme.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Wireless sensor networks (WSNs) [Ver+10] have attracted much interests in the last decade. They are composed by autonomous nodes, with sensing capability of some physical phenomenon (*e.g.* temperature, pressure, or radioactivity level). Their application area ranges from environmental or traffic monitoring to industrial domain. This thesis addresses several problems related to the efficient information collection as well as distributed estimation in a WSN.

1.1.1 Network coding for distributed data collection

In order to ensure ease of deployment and robustness, the communication between the nodes might need to be performed in absence of designated access points and of a hierarchical structure. At the network layer, dissemination of the measurements to all the nodes can be achieved using an asynchronous protocol based on *network coding* (NC) [Ahl+00] in order to reduce the traffic. In the protocol, each node in the network broadcasts a packet evaluated as the linear combination of the local measurement, and of the packets received from neighboring nodes. The linear coefficients are randomly chosen, and are sent in each packet header. Upon an appropriate number of communication rounds, each node has collected enough linearly independent combinations of the network measurements, and can perform decoding, by solving a system of linear equations.

Without considering the correlation of the source, the number of linearly independent combinations should be equal to the number of sensor nodes in the network. In many practical situations, the measurements obtained by the WSN are spatially and temporally correlated. This correlation can be exploited to perform decoding, as done in [IKAA12; Bas+12; RAG12; BTF12]. The first objective of this thesis is to investigate how many linearly independent combinations are necessary and sufficient to recover the source information.

1.1.2 Distributed self-rating

The numerous nodes in a WSN may have different sensing abilities and each node may only have a partial knowledge of the characteristics of the noise of its sensors, *e.g.*, biased or unbiased as in [Chi+11]. It may thus be interested in getting an estimate of these characteristics. Similarly, in crowdsensing applications [Guo+15], the reliability of the sensing services (SaaS) [She+13] requires high quality measurements that are not always available at all devices, since the sensing accuracy

is highly dependent on the embedded sensors. This problem is addressed, *e.g.*, by introducing reputation-based mechanisms [YS12; Han+13; KM14a; KM14b; Ren+15]. In such schemes data are collected from smartphones, based on an auction mechanism. The sensing tasks are assigned by some central authority to the agents on the basis of their reputation level, to maximize the utility for the crowd. Such reputation-based mechanisms are effective in rating agents, but need data centralization and some rating authority. Allowing agents to assess their ability has the benefit to help them providing measurements as well as their associated quality levels. This facilitates further measurement processing and may avoid compromising the reputation of an agent temporarily producing outliers. Moreover, in SaaS application, agents knowing that their sensing ability is above the average may negotiate a better reward for their measurements.

Assume that each node or agent does not know initially the characteristics of its sensor measurements and it is willing to self-rate these characteristics. The rest of this thesis is to design and analyze some *distributed self-rating*(DSR) algorithm allowing each node in a network to self-rate the characteristics of its sensors. More generally, the proposed approaches allow the agents of a network to self-rate their level of ability (LoA) at doing some task (sensing, detection, classification...).

We start with a sub-problem of DSR, namely *distributed faulty node detection* (DFD), by considering only two levels of the performance of sensors, *i.e.*, good and defective. Defective sensors are defined as those sensors who frequently produce *outliers*, *i.e.*, abnormal measurements which cannot be justified by the mere effects of sensing noise [BL94]. The presence of nodes equipped with defective sensors may significantly altered the behavior of a WSN dedicated to distributed estimation task. It is thus very important to detect such nodes efficiently (with low communication and computation costs), ideally in a distributed way.

1.1.3 Considering an intermittent connectivity

The delay/disruption-tolerant networks (DTNs) are challenging networks with dynamic topology and sparse connectivity. The communications between nodes of a DTN are quite limited and have a short duration. Due to their intermittent connectivity, inference and learning over DTNs is much more complicated than in traditional networks, see, *e.g.*, [HCY11; SRF14; Zhu+14; GLG16; Pan+15].

For a large WSN with extreme power constraint, in order to conserve power, nodes should not always communicate with the others. That is why a DTN architecture is important in the design of a WSN in many application scenarios [Fal03; KAF12].

In the last part of the thesis, DFD algorithms are proposed for DTNs, as a following of DFD in dense WSNs. The results are then generalized to the DSR problem by considering a finite number of possible LoA. Nodes are assumed to have only sporadic pairwise interactions, as is typically the case in DTNs, or in networks where exchanges are performed via gossiping [FZ08; Dim+10].

1.2 State of the Art and Main Contributions

This section presents the related work and the contribution of this thesis in different topics.

1.2.1 Network coding for data compression in WSNs

Joint source-network coding in WSNs with correlated source data can be performed using techniques borrowed from compressive sensing (CS) [Don06]. The well-known CS problem refers to the compression of a vector $\boldsymbol{\theta} \in \mathbb{R}^N$, obtained by acquiring linear measurements whose number M can be significantly smaller than N. If $\boldsymbol{\theta}$ is k-sparse with respect to some known basis, its almost surely exact reconstruction can be evaluated from the linear measurements using basis pursuit, for M as small as $\mathcal{O}(k \log(N))$ [CRT06; CT06a]. The same result holds true also for compressible vector $\boldsymbol{\theta}$ [Don06], with reconstruction quality matching the one allowed by direct observation of the biggest k coefficients of $\boldsymbol{\theta}$ in the transform domain. The major feature of compressed sensing is that the linear coefficients do not need to be adaptive with respect to the signal to be acquired, but can actually be random, provided that appropriate conditions on the global measurement matrix are satisfied [CT06a; CT05]. Moreover, compressed sensing is robust to the presence of noise in the measurements [CT05; HN06].

Bayesian compressed sensing [JXC08] refers to the same problem, considered in the statistical inference perspective. In particular, the vector to be compressed is now understood as a statistical source $\boldsymbol{\Theta}$, whose *a priori* distribution can induce sparsity or correlation between the symbols. In practical implementations, estimation from the linear measurements can be achieved exploiting statistical graphical models [Mon12], *e.g.*, using belief propagation as done in [BSB10] for deterministic measurement matrices, and in [BM11] for random measurement matrices.

The main difficulty of the direct application of CS in WSNs comes from the fact that usually, sensor readings are real-valued or quantized while NC operations are done in finite fields. As a consequence, the NC matrix, which plays the role of the sensing matrix in CS, has elements in a finite field, which makes it difficult to exploit the efficient reconstruction algorithms of CS. A first solution to this problem is proposed in [FM11], where NC is performed in the real field. This framework is also applied in [LCJ15], where data reconstruction is performed progressively to reduce the decoding delay. In [NL14], linear combinations resulting from real-field NC are quantized before transmission. The price to be paid by all these technique is larger headers and an incompatibility with classical NC.

We consider random linear network coding (RLNC) [Ho+06], where both the quantization levels for the measurements and the NC coefficients are chosen in the same finite field. With RLNC, the coding coefficients are randomly chosen and sent in each packet header [CWJ03]. Usually, NC vectors are sparse: packets reaching the sink are seldom combinations of all packets and depend on the network topology. This allows compressing efficiently the NC vectors [Jaf+09] and reduces the decoding complexity [Bas+12].

The considered problem can be formulated as the estimation of a source vector $\boldsymbol{\theta}$ (of N symbols), obtained by acquiring M non-adaptive linear measurements, with $M < N$. In a Bayesian setting, $\boldsymbol{\theta}$ is understood as a realization of a random source $\boldsymbol{\Theta}$, whose *a priori* distribution induces compressibility (sparsity and/or correlation between the symbols). The reconstruction problem is formulated in this case as an estimation problem, solvable using standard Bayesian techniques, *e.g.*, Maximum *A Posteriori* (MAP) estimation. This setting has been considered in [IKAA12], where exact MAP reconstruction is obtained solving a mixed-integer quadratic program, and in

[Bas+12; BTF12; RAG12], where approximate MAP estimation is obtained using variants of the belief propagation algorithm [KFL01].

This problem can also be defined as Bayesian compressed sensing over finite fields, of which the performance analysis has been addressed in [DM09; RAG12], and [SL13]. The work in [SL13] does not consider Bayesian priors, and assumes a known sparsity level of $\boldsymbol{\theta}$. Ideal decoding via ℓ_0-norm minimization is assumed, and necessary and sufficient conditions for exact recovery are derived as functions of the size of the vector, its sparsity level, the number of measurements, and the sparsity of the sensing matrix. Numerical results show that the necessary and sufficient conditions coincide, as the size of $\boldsymbol{\theta}$ asymptotically increases. A Bayesian setting is considered in [DM09] and [RAG12]. In [DM09] a prior distribution induces sparsity on the realization of $\boldsymbol{\Theta}$, whose elements are assumed statistically independent. Using the method of types [Csi98], the error exponent with respect to exact reconstruction using ℓ_0-norm minimization is derived in absence of noise in the measurements, and the error exponent with respect to exact reconstruction using minimum-empirical entropy decoding is derived for noisy measurements. In [RAG12] specific correlation patterns (pairwise, cluster) between the elements of $\boldsymbol{\Theta}$ are considered. Error exponents under MAP decoding are derived, only in case of absence of noise on the measurements.

This work investigates the compression efficiency of RLNC. The contribution of this work can be summarized as follows. We assume a Bayesian setting and we consider MAP decoding. Inspired by the work in [SL13], we aim to derive necessary and sufficient conditions for almost surely exact recovery of $\boldsymbol{\theta}$, as its size asymptotically increases. We consider

- a sparse coding matrix with entries of different sparsity level, which is a more realistic hypothesis and no analysis has been previously performed for such model.
- two kinds of noise: *a*) the sensing noise, affecting the measurements prior to network coding (*i.e.,* prior to random projection acquisition in the CS framework); *b*) the additive noise, affecting the network coded packets (*i.e.*, the random projections in the cs framework). To the best of our knowledge, no analysis has been previously performed in presence of both kinds of noise.
- three classes of prior distributions on the source vector: *i*) the prior distribution is sparsity inducing, and the elements are statistically independent; *ii*) the vector $\boldsymbol{\Theta}$ is a Markov process; *iii*) the vector $\boldsymbol{\Theta}$ is an ergodic process. To the best of our knowledge, no analysis has been previously performed for the latter source model, which is quite general.

Considering source model *i*), our results for the noiseless setting are compatible with the ones presented in [SL13]; in addition, we can formally prove the asymptotic convergence of necessary and sufficient conditions, and extend the bounds on the sparsity factor of the sensing matrix in presence of additive noise. The asymptotic analysis under MAP decoding, both for the noiseless case and in presence of additive noise *b*), are compatible with the results derived in [DM09], respectively under ℓ_0-norm minimization decoding and under minimum-empirical entropy decoding. Error exponents for MAP decoding of correlated sources in the noiseless setting are compatible with the ones presented in [RAG12], and are here extended to the case of arbitrary statistical structure, and presence of noise contamination both preceding and following the sensing operation.

1.2.2 Distributed faulty node detection in WSNs

Many efforts have been considered in the detection of nodes with defective sensors in WSNs, see [ZMH10; MK13; Don+14] and references therein. As far as the DFD is concerned, in particular, many different techniques have been proposed in the literature (see the survey paper [MK13] for a detailed classification). Some techniques rely only on data coming from the direct neighborhood of each node [CKS06], others also use data coming for farther nodes [DP+15; Bra+13]. Some techniques are model-based and rely on a parametric model of the system observed by the sensors to identify outliers [Ise05; Lo+13], whereas model-free techniques mainly exploit statistical properties of the measurements, such as spatial and temporal correlation. For example, the algorithm in [CKS06] allows each node to estimate its own status by comparing its local measurements with those at neighboring nodes. In a first phase only a tendency status (*good*, *faulty*, *likely good*, or *likely faulty*) is determined. In the second phase the tendency status of neighbors are collected and associated to their measurements to obtain a more reliable assessment. In [Din+05] the local test is based on the comparison between the local measurement and the median of the measurements of the neighbors. In [PK15], a modified three-sigma edit test is proposed to identify nodes producing measurements with very high variance. Iterative algorithms are proposed in [GXL07; Ji+10], where the weighted-median and the weighted average criterion are considered, respectively. In both cases the local test weights the measurements of the neighbors by the confidence level obtained from the previous detection rounds, under the assumption of permanent node failure. In [LC08], the algorithm uses time redundancy to tolerate transient faults in sensing and in communication. The adaptive algorithm proposed in [Cho+09] adjusts the decision threshold at each round to improve its detection accuracy. In [Bra+13], a generic outlier identification function is assumed to be available at each node. The results provided by this function at each node as well as measurements are exchanged to allow the whole network to identify all outliers in a given dataset. In [DP+15], a distributed Bayesian approach is proposed to detect the outliers in a large set of data collected by a WSN.

In general the availability of more data at a given node facilitates the detection and identification of outliers, at the price of higher communication costs and detection delays. An interesting approach to reduce the number of tests is group testing (GT), which is a statistical method allowing to identify a small number of outliers within a large set of data by performing only a limited number of elementary tests, referred to as LODTs, on data subsets [DH93]. The elementary test is only able to determine whether outliers are present in the considered set. A distributed GT-based algorithm extending the results of centralized GT [Che+11] is proposed in [TTF13], under the assumption of vanishing ratio of defective sensors. The solution provided in [TTF13] suffers however from a significant drawback in the communication overhead necessary for node indexing and for the dissemination of the test results from the cluster heads to the rest of the wireless sensor network (WSN).

Despite the significant number of papers on this topic, there are still open issues to be addressed, as pointed out in [MK13; Don+14]. It is well known, in this regard, that WSNs are usually constrained in terms of memory, computational power and energy sources. Most of the traditional fault detection schemes, conceived for wired networks, have paid little attention to these issues. They are

usually memory and energy eager, and are often high demanding in terms of communication effort and computational complexity. It is certainly true, on the one hand, that the exchange of messages and their processing are the only means of fault diagnosis; on the other hand, however, outlier detection schemes targeted to WSNs must be lightweight and impose a limited communication cost.

The literature is mainly focused on new or improved LODTs, which are validated considering specific case studies through simulation, thus without determining *a priori* the relationship between the parameters of the algorithm (*e.g.*, detection threshold) and the performance. At the our best knowledge, no attempts are present to widen the perspective, generalizing the DFD methodology and providing equilibrium or stability analysis of iterative DFD algorithms. The latter are very important issues as the knowledge of the conditions under which the algorithm reaches the equilibrium and is stable, prevents the system to fail in an unpredictable way.

In Chapter 3, we propose two DFD algorithms inspired by the GT technique: a single-decision one and an iterative one, the second being better suited to WSNs with a large proportion of defective sensors. In both cases we decouple the LODT, that provides only an indication (true, false) of the existence of outliers in the neighborhood, from the final decision on the node status (good, faulty). The proposed algorithms consist, in fact, of two stages, namely *local outlier detection* using a LODT, and *decision*. In the first stage, each node collects data from its neighborhood, performs the LODT to determine whether outliers are present among the collected data, and broadcasts the result. Then, each node estimates the status of its sensor in the decision stage, on the basis of the outcomes of the LODTs performed in the neighborhood.

Sufficient conditions to be satisfied by the LODT have been established to ensure the existence of an equilibrium of the DFD algorithms. These results are general, since no specific form of the LODT need to be considered. We show that the LODT is only required to determine whether outliers are present in a set of data, without necessarily being able to identify which data are erroneous, thus widening the range of applicable LODTs to those requiring a few measurements from neighboring nodes. Such LODTs are easily accessible, *e.g.*, in the context of bounded-error estimation [Mil+96; Jau+01] and may provide good results even with as few as two or three measurements available to perform the test. This aspect is crucial in the WSNs context because it allows to relieve the network traffic and prolong its lifetime. Let us stress, in this regard, that most of traditional outlier detection tests (based on the median [GXL07] or the mean [Ji+10] of measurements, for instance) are not very efficient when only few measurements are available. Furthermore, the low complexity of such LODTs is well-suited to the stringent resource (memory, computational power, etc.) constraints that usually affect sensor nodes.

Finally, the trade-off between false alarm probability and detection probability is characterized theoretically and by simulation for some outlier models in realistic channel conditions.

Summarizing, the main novelties of this work are:

- Proposed a low-complexity DFD algorithm based on very generic LODTs that have the potential to work with very few measurements;
- Found the mild conditions to be satisfied by the LODT and by the density of nodes in the network to ensure that the iterative DFD algorithm leads to an equilibrium;

- Characterized analytically the performance in realistic channel conditions, related to packet losses and to channel access issues at the MAC layer, the latter being seldom considered in other papers;
- Implementation of the proposed algorithm on a real testbed to verify the effectiveness of the algorithm.

1.2.3 Distributed faulty node detection in DTNs

The DFD problem is much less investigated in DTNs. When only few measurements are available, classical LODTs are not very efficient: for example, they cannot, based on two measurements only, determine which node is defective. This is a typical situation in DTNs when two nodes meet, take measurements, and share these measurements. Applying directly classical DFD algorithms in DTNs may thus be quite ineffective.

A closely related problem has been previously considered in [Zhu+12] in the context of Vehicular DTNs (VDTNs). A large number of sensor nodes are fixed and some vehicles, called mobile carriers (MC) collect data from these sensors. The sensor nodes can only communicate with the MCs in their vicinity. A MC needs to collect enough measurements to perform a test to decide which have been produced by defective sensors. Once a defective node is deemed defective by a MC, it is added to its blacklist. The MC provides information to sensors about their status. MCs also exchange their blacklists to accelerate the faulty node detection.

In [Pen+14], a similar problem of distributed malware detection in DTN is addressed. Each node evaluates after the meeting with another node whether the latter has performed suspicious actions (malware transmission trial). When after several meetings of Node j, Node i detects often suspicious activities, a cut-off decision is performed against Node j, which is ignored in next meetings. The main drawback of this approach is the long time required to identify and isolate misbehaving nodes. Misbehavior detection in DTN is also considered in [Zhu+14; HO+15], where the DTN is perturbed by routing misbehavior caused by selfish or malicious nodes. The identification approach in [Zhu+14] is not distributed, since a Trusted Authority periodically checks the forwarding history of nodes to identify possible misbehavior. A collaborative approach is proposed in [HO+15], where each node can detect whether the encountered node is selfish using a local watchdog. The detection result is disseminated over the network to increase the detection precision and to reduce the delay. Trust/Reputation management is another important aspect to help DTNs resist various potential threats. For example, [AF12] provides an iterative trust management mechanism to fight against Byzantine attacks in which several nodes are totally controlled by the adversary. In [Liu+15], a defense against Sybil attacks [Dou02] is introduced, which is based on the physical feature of the wireless propagation channel. A trust model for underwater acoustic sensor networks is presented in [Han+15] to take into account several trust metrics such as link trust, data trust, and node trust.

In Chapter 4, we present a fully distributed and easily implementable algorithm to allow each node of a DTN to determine whether its own sensors are defective. Due to the communication conditions of DTNs, DFD algorithm in DTNs is quite different from those in WSNs, despite of their same settings of LODT. The analysis of the properties of the algorithm is also totally different. This

work shows that the behavior of the proposed DFD algorithm can be analyzed using Markov models and tools borrowed from control theory and population dynamics. For that purpose, the belief of each node about the status of its sensors is quantized. The evolution of these quantized beliefs are then shown to follow two Markov chains. The dynamic of the proportions of nodes with a given belief is then analyzed. Sufficient conditions on the decision parameters to ensure the existence and unicity of an equilibrium of the DFD algorithm are then provided. Given the characteristics of the LODT, upper and lower bounds of the *detection rate*, *i.e.*, proportion of nodes which have effectively identified their sensors as defective, and of the *false alarm rate*, *i.e.*, proportion of nodes which believe that their good sensors are in fact defective, are also obtained. These theoretical results provide guidelines to properly choose the parameters of the algorithm.

1.2.4 Distributed Self-Rating by Pairwise Interaction

Consider a community of agents collaborating to execute some task (sensing, detection, classification, see, *e.g.*, [Luo+07; Ang+09; SBW16]), for which it is reasonable to expect agents showing different levels of ability (LoA). This ability can be related, for example, to the quality of the sensors they are equipped with. Initially, each agent does not know how well it performs in comparison with its peers and it is thus willing to assess its own ability.

In Chapter 5, one assumes that agents are able to have only sporadic pairwise interactions, as in DTNs, or in networks where exchanges are performed via gossiping [FZ08; Dim+10]. Based on these assumptions, a distributed algorithm allowing each agent to assess its own LoA at doing some task is presented. This individual assessment procedure involves pairwise interactions with the peers and local comparison tests (LCT), providing an estimate, among two interacting agents, of the better-performing one. Similar LCT have been considered in [Hec+16], where a central authority aims at ranking agents based on the results of LCTs. The central authority has the power to select the interacting pairs of agents and the sequence of interactions to obtain the fastest ranking.

Previous instances of the DSR problem have been considered in [Chi+11; FFR14], where each node of a WSN has to estimate the value of some constant parameter from noisy scalar measurements, as well as the bias [Chi+11] or the level of the variance [FFR14] of its measurement noise, both assumed to take only two possible values. The nodes are partitioned into two classes in [Chi+11], each class being characterized by the absence or presence of measurement bias. The algorithm involves a gossip consensus, robust against node mobility, and a distributed ranking of the agents [FZ08] according to their measured value. Each node is then able to decide whether its measurement noise is biased. The partitions considered in [FFR14] depend on the possible values of the measurement noise variance. Cooperation is achieved via consensus algorithms [HM09]. The LOA of a node is determined just by comparing the local measurement with the estimate of the common parameter.

The DSR problem can be viewed as a generalization of DFD, in which the number of LOA is 2. When the number of LoA is equal to the number of agents, the DSR problem is equivalent to the distributed self-ranking problem addressed in a centralized way in [Hec+16] and with a distributed approach in [FZ08].

DSR can also be viewed as a type of distributed classification problem [Luo+07; Ang+09]. Nevertheless, the latter is aimed at classifying *data* with network of agents, whereas DSR in this paper is to make each agent estimate its *own* LoA. Agents aim thus at performing a self-classification.

Chapter 5 extends the DTN-DFD to DSR considering a finite number of possible LoA for each agent. One assumes here that there is an agreement on the proportions of agents with similar LoA, and that each agent knows the characteristics of the LCT. Then, the proposed DSR algorithm involves only the results of LCTs performed when two agents interact. One shows that each agent, based on the proportion of interactions during which it was deemed better than the other agents, is able to iteratively estimate its LoA. This work extends also previous results obtained in the context of DSR by [Chi+11; FFR14] in several directions. The DSR algorithm involves a generic LCT, only characterized by its probabilities of error. Many situations can be described with such LCT, ranging from the noisy measurement with possible bias of some constant parameter as in [Chi+11], to the result of supervised image classification tests performed by pairs of agents, via the results of blitz-games when agents have to assess their level in a game.

The effectiveness of the proposed algorithm is measured by the *correct decision rate* (CDR) and by the *false decision rate* (FDR), *i.e.*, the proportion of agents that correctly and erroneously self-rate their level of expertise. The analysis of the DSR algorithm is performed assuming a well-mixed population of agents with intercontact delay following an exponential distribution [Zhu+10; GLG16; HO+15; Pan+15]. This is somewhat more restrictive than the connected communication graph model with sporadic communication considered in [Chi+11]. The communication model considered here allows describing the evolution with time of the proportions of agents with similar beliefs in their LOA by continuous-time state equations. The existence of an equilibrium is shown. Closed-form expressions for the various proportions of agents with similar beliefs in their LOA is provided at equilibrium. The dependence of the CDR and FDR in the characteristics of the LCT provides insights on the way the DSR algorithms should be tuned to get a compromise between CDR and FDR. Simulation results in the context of agents equipped with sensors aiming at determining the performance of their sensors show and excellent match with theoretical results.

1.3 Publications

Journal Papers The following is a list of journals papers produced during my Ph.D. candidature.

1. W. Li, F. Bassi, and M. Kieffer, "Robust Bayesian compressed sensing over finite fields: asymptotic performance analysis", in preparation. [LBK14]

2. W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests", *IEEE transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp 29-48, 2016. [Li+16a]

3. W. Li, F. Bassi, L. Galluccio, and M. Kieffer,"Distributed Faulty Node Detection in Delay Tolerant Networks: Design and Analysis", *IEEE transactions on Mobile Computing*, submitted. [Li+16d]

4. W. Li, F. Bassi, L. Galluccio, and M. Kieffer, "Peer-assisted Individual Assessment in a Multi-agent system", *Automatica*, submitted. [Li+16f]

Conference Papers The following is a list of journals papers produced during my Ph.D. candidature.

1. W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Low-Complexity Distributed Fault Detection for Wireless Sensor Networks", *Proc. IEEE International Conference on Communications, ICC 2015*, London, UK, 2015, pp. 6712–6718. [Li+15b]

2. W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Iterative Distributed Outlier Detection for Wireless Sensor Networks: Equilibrium and Convergence Analysis", *Proc. IEEE Conference on Decision and Control, CDC 2015*, Osaka, Japan, 2015, pp. 3050-3056. [Li+15a]

3. W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Impact of Channel Access Issues and Packet Losses on Distributed Outlier Detection within Wireless Sensor Networks", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2016*, Shanghai, China, 2016. [Li+16c]

4. W. Li, F. Bassi, and M. Kieffer, "Sparse Random Linear Network Coding for Data Compression in WSNs", *Proc. IEEE International Symposium on Information Theory, ISIT 2016*, Barcelona, Spain, 2016, to appear. [LBK16]

5. W. Li, L. Galluccio, M. Kieffer, and F. Bassi, "Distributed Faulty Node Detection in DTNs", *Proc. International Conference on Computer Communication and Networks, ICCCN 2016*, Hawaii, USA, 2016, to appear. [Li+16b]

6. W. Li, F. Bassi, L. Galluccio, and M. Kieffer, "Self-Rating in a Community of Peers", *IEEE Conference on Decision and Control, CDC 2016*, Las Vegas, USA, 2016, to appear. [Li+16e]

Chapter 2

Sparse Random Linear Network Coding for Data Compression in WSNs

This chapter contains an analysis of some information gathering problems in the WSNs. The random linear network coding is considered for data compression, and an information-theoretic approach is applied to demonstrate the necessary and sufficient conditions to realize the asymptotically perfect reconstruction under MAP estimation. Several related work has been discussed in Section 1.2.1.

The chapter is organized as follows. Section 2.1 presents a motivation example and shows that in practice the sparsity factors of different entries in a network coding matrix are usually different. Section 2.2 introduces the considered signal models in the context of data dissemination in a wireless sensor network. In Section 2.3, we derive the necessary conditions for asymptotic almost surely exact recovery, both for the noiseless and noisy cases. Section 2.4 describes the sufficient conditions and the error exponents under MAP decoding, for the noiseless case and in presence of additive noise only. In Section 2.5, sensing noise is also taken into account. Section 2.6 concludes the chapter.

2.1 Motivating Example

Consider the WSN whose topology is shown in Figure 2.1(a). $S = 29$ wireless sensor nodes are uniformly distributed over a square of unit area with a sink node located at the center. Assume that each sensor Node i attempts to transmit a binary measurement θ_i to the sink using RLNC over the Galois field $\mathbb{F}_2$. Time is slotted and each sensor broadcasts a packet or combination of packets in its time slot, starting from the nodes the farthest from the sink. Each node can only communicate with its neighbor nodes at a distance less than 0.25. When Node i performs a combination of packets, its local measurement θ_i is assumed to be always used, whereas a packet received from one of its neighbors has a probability ν to be involved. After N_{R} communication rounds, the sink is able to build an NC matrix with $M = N_{\mathrm{R}} S_0$ lines and S columns, where S_0 is its number of one-hop neighbors. The sparsity of each column of the NC matrix is then evaluated. Figure 2.1(b) represents the proportion of non-zero coefficients of each column parametrized in ν averaged over 10^6 independent Monte-Carlo simulations of the network. Only connected networks are considered in the simulations. The average value of M is $17 < S$. Columns associated to nodes which are far from the sink are sparser than columns associated to nodes close to the sink. Moreover, a small value of ν makes the overall coding matrix sparser. This justifies our study considering coding matrices with sparsity varying among columns.

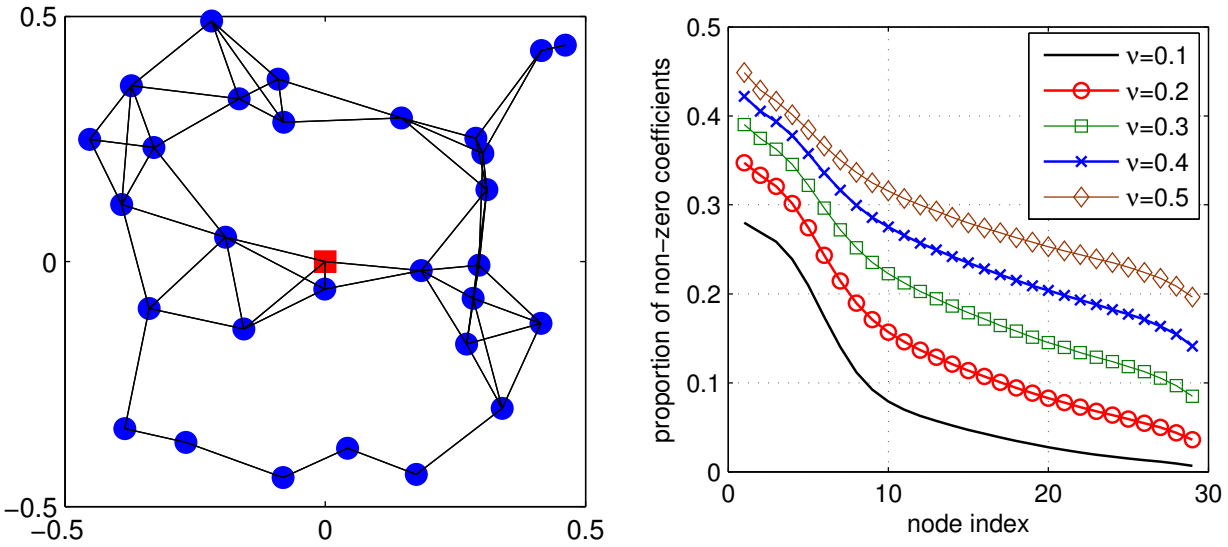

FIGURE 2.1: (a) example network topology ; (b) evolution of the proportion of non-zero NC coefficient as a function of the node index sorted by increasing distance to the sink, for $\nu \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$.

2.2 System Model and Problem Setup

This section introduces the system model as well as various hypotheses on the sources and on the sensing and additive noises. In what follows, sans-serif font denotes random quantities while serif font denotes deterministic quantities. Matrices are in bold-face upper-case letters. A length n vector is in bold-face lower-case with a superscript n. Calligraphic font denotes set, except $\mathcal{H}$, which denotes the entropy rate. All logarithms are in base 2.

2.2.1 The source model

Consider a wireless sensor network consisting of a set $\mathcal{N}$ of $N = |\mathcal{N}|$ sensors. The target physical phenomenon (*e.g.* the temperature) at the n-th sensor is represented by the random variable $\mathsf{\Theta}_n$, taking values on a finite field $\mathbb{F}_Q$ of size Q. Let $\boldsymbol{\theta}^N$ be a realization of the random vector $\mathsf{\Theta}^N = (\mathsf{\Theta}_1, \ldots, \mathsf{\Theta}_N)$, taking values in $\mathbb{F}_Q^N$. The vector $\mathsf{\Theta}^N$ represents the source in the Bayesian compressed sensing framework. The probability mass function (pmf) associated with $\mathsf{\Theta}^N$ is denoted by $p\left(\boldsymbol{\theta}^N\right)$, rather than $p_{\mathsf{\Theta}^N}\left(\boldsymbol{\theta}^N\right)$, for the sake of simplicity. In general, the analytic form of $p\left(\boldsymbol{\theta}^N\right)$ depends on the characteristics of the observed phenomenon and on the topology of the sensor network. Here we consider three different models, defined as follows.

SI: Sparse, Independent and identically distributed source. Each element of the source vector $\mathsf{\Theta}^N$ is independent and identically distributed (iid) with pmf $p_{\mathsf{\Theta}}(\cdot)$ and $p_{\mathsf{\Theta}}(0) > 0.5$,

$$p\left(\boldsymbol{\theta}^N\right) = \prod_{n=1}^{N} p_{\mathsf{\Theta}}\left(\theta_n\right). \tag{2.1}$$

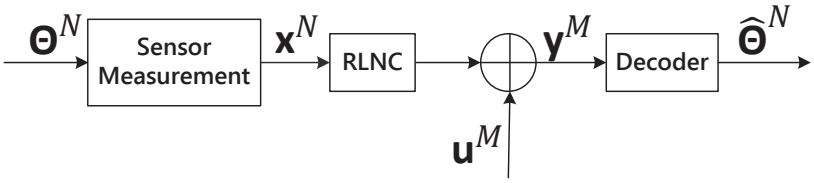

FIGURE 2.2: Block diagram for network compressive sensing model

StM: Stationary Markov model. Let $\boldsymbol{\theta}_n^{n+r-1} \in \mathbb{F}_Q^r$ denote the sequences $(\theta_n, \ldots, \theta_{n+r-1})$. This is the stationary r-th order Markov model with $r \in \mathbb{N}^+$ and $1 \leqslant r \ll N$ and transition probability $p\left(\theta_{n+r} \mid \boldsymbol{\theta}_n^{n+r-1}\right)$. The pmf of $\boldsymbol{\Theta}^N$ may be written as

$$p\left(\boldsymbol{\theta}^N\right) = p\left(\boldsymbol{\theta}_1^r\right) \prod_{n=1}^{N-r} p\left(\theta_{n+r} \mid \boldsymbol{\theta}_n^{n+r-1}\right). \tag{2.2}$$

GSE: General Stationary and Ergodic model. This is the general case, without any further assumption apart from the ergodicity of the source.

2.2.2 The sensing model

The considered system model is shown in Figure 2.2.

Let $\mathsf{x}_n \in \mathbb{F}_Q$ be the measurement of Θ_n obtained by the n-th sensor. The random vector $\mathbf{x}^N = (\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_N) \in \mathbb{F}_Q^N$ is a copy of the source vector $\boldsymbol{\Theta}^N$ corrupted by the *sensing noise*. The sensing noise models the effect of imperfect measure acquisition at each sensor. It is described by the stationary transition probability $p_{\mathsf{x}|\Theta}(x_n \mid \theta_n)$, $\forall n$. Remark that this implies that $\mathbf{x}^N$ is stationary as long as $\boldsymbol{\Theta}^N$ is stationary. The local measurement x_n at node n is used to compute a packet via RLNC [Ho+06], which is then broadcast and received by the neighbors of n. Each node in the network can act as a sink, and attempt reconstruction of $\boldsymbol{\Theta}^N$, after a number $M \leqslant N$ of linear combinations has been received. The effects of RLNC at a sink node can be modeled as multiplying $\mathbf{x}^N$ by a random matrix $\mathbf{A} \in \mathbb{F}_Q^{M \times N}$. We assume that some *additive noise* $\mathbf{u}^M \in \mathbb{F}_Q^M$ affects the received packets, modeling the effects of transmission. Each entry of $\mathbf{u}^M$ is iid with pmf $p_{\mathsf{u}}(\cdot)$. The sink node is assumed to have received M packets, with the i-th packet carrying the coefficients $\mathbf{A}_i$ and the result of linear combination $\mathsf{y}_i \in \mathbb{F}_Q$, where $\mathbf{A}_i$ is the i-th row of $\mathbf{A}$ and $\mathsf{y}_i = \mathbf{A}_i \mathbf{x}^N + \mathsf{u}_i$, with all operations in $\mathbb{F}_Q$. The vector $\mathbf{y}^M = (\mathsf{y}_1, \mathsf{y}_2, \ldots, \mathsf{y}_M)^{\mathrm{t}} \in \mathbb{F}_Q^M$ can be then represented as

$$\mathbf{y}^M = \mathbf{A}\mathbf{x}^N + \mathbf{u}^M, \tag{2.3}$$

where the network coding matrix $\mathbf{A}$ plays the role of the random sensing matrix in the compressed sensing setup. Depending on the presence of the sensing and additive noises, one obtains four types of noise models, namely Without Noise (WN), Noise in Communications (NC) only, Noise in the Sensing process (NS) only, and noise in both Communications and in the Sensing process (NCS). These models are summarized in Table 2.1.

In general, the matrix $\mathbf{A}$ is not necessarily of full rank, and it is assumed to be independent on $\mathbf{x}^N$. Two different assumptions about the structure of $\mathbf{A}$ are considered here: ($A1$) the entries of $\mathbf{A}$

TABLE 2.1: Classification and notation based on the presence of noise

		Additive Noise	
		absent	present
Sensing Noise	absent	WN	NC
	present	NS	NCS

are iid, uniformly distributed in $\mathbb{F}_Q$; *(A2)* the entries of $\mathbf{A}_i$, the i-th column of $\mathbf{A}$, are assumed to be iid in $\mathbb{F}_Q$ with pmf

$$p_{\mathsf{A}_i}(0) = 1 - \gamma_i, \qquad p_{\mathsf{A}_i}(q) = \frac{\gamma_i}{Q-1} \text{ for } q \in \mathbb{F}_Q \setminus \{0\} \tag{2.4}$$

where γ_i is the sparsity factor of $\mathbf{A}_i$. The sparsity factors γ_i and γ_j of two different columns $\mathbf{A}_i$ and $\mathbf{A}_j$ may be different. Define $\gamma = \min_{i=1,\dots,N} \gamma_i$ as the minimum sparsity factor. We only assume that

$$0 < \gamma \leqslant 1 - Q^{-1}. \tag{2.5}$$

Notice that choosing $\gamma < 1 - Q^{-1}$ corresponds to assumption *(A2)*, while choosing $\gamma_1 = \dots \gamma_N = 1 - Q^{-1}$ corresponds to assumption *(A1)*, since (2.4) becomes the uniform distribution.

In practice, sparser matrices are preferable. As the information of the network coding matrix is carried in the headers of packets [CWJ03; Jaf+09], the network coding overhead may be large if $\mathbf{A}$ is dense and N is large. Moreover, as mentioned in [Bas+12], sparse matrices facilitate the convergence of the approximate belief propagation algorithm [KFL01]. Notice that, in practice, the structure of $\mathbf{A}$ is strongly dependent on the structure of the network. For example, [RAG12] assumes that only a subset of sensors $\mathcal{S}_i \subset \mathcal{N}$ have participated in the i-th linear mixing. The content of the subsets $\mathcal{S}_i$ depends on the location of each sensor and is designed to minimize communication costs. In $\mathbf{A}$, coefficients associated to nodes belonging to $\mathcal{S}_i$ follow a uniform distribution, while the others are null. The model in [RAG12] is not considered here, since we aim at a general asymptotic analysis, independent on the topology of the network.

2.2.3 MAP Decoding

The sink node observes the realization $\mathbf{y}^M$ and perfectly knows the realization $\mathbf{A}$, *e.g.*, from packet headers, see [CWJ03] and [Jaf+09]. The maximum *a posteriori* estimate $\hat{\boldsymbol{\theta}}^N$ of the realization of $\boldsymbol{\Theta}^N$ is evaluated as

$$\hat{\boldsymbol{\theta}}^N = \arg \max_{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N \mid \mathbf{y}^M, \mathbf{A}\right), \tag{2.6}$$

where the *a posteriori* pmf is

$$\begin{aligned} p\left(\boldsymbol{\theta}^N \mid \mathbf{y}^M, \mathbf{A}\right) &\propto p\left(\boldsymbol{\theta}^N, \mathbf{y}^M, \mathbf{A}\right) \\ &= \sum_{\mathbf{x}^N \in \mathbb{F}_Q^N} \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N, \mathbf{x}^N, \mathbf{u}^M, \mathbf{y}^M, \mathbf{A}\right) \\ &= \sum_{\mathbf{x}^N \in \mathbb{F}_Q^N} \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{x}^N \mid \boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) p\left(\mathbf{A}\right) p\left(\mathbf{y}^M \mid \mathbf{x}^N, \mathbf{u}^M, \mathbf{A}\right). \end{aligned} \tag{2.7}$$

Note that the conditional pmf $p\left(\mathbf{y}^M \mid \mathbf{x}^N, \mathbf{u}^M, \mathbf{A}\right)$ is an indicator function, *i.e.*,

$$p\left(\mathbf{y}^M \mid \mathbf{x}^N, \mathbf{u}^M, \mathbf{A}\right) = 1_{\mathbf{y}^M = \mathbf{A}\mathbf{x}^N + \mathbf{u}^M}. \tag{2.8}$$

An error event (decoding error) occurs when $\hat{\boldsymbol{\theta}}^N \neq \boldsymbol{\theta}^N$, with probability

$$P_{\mathrm{e}} = \Pr\left\{\hat{\boldsymbol{\Theta}}^N \neq \boldsymbol{\Theta}^N\right\}. \tag{2.9}$$

Our objective is to evaluate lower and upper bounds of (2.9) under MAP decoding, as functions of M, N, and γ, for the various source and noise models previously introduced. With these bounds, one can obtain necessary and sufficient conditions on the ratio M/N for asymptotic (with $N \to \infty$) perfect recovery, *i.e.*, to obtain

$$\lim_{N \to \infty} P_{\mathrm{e}} = 0. \tag{2.10}$$

2.3 Necessary Condition for Asymptotic Perfect Recovery

This section derives the necessary conditions for asymptotically ($N \to \infty$) vanishing probability of decoding error. They only depend on the assumptions considered about the sensing and additive noises. We directly analyze the NCS case for the GSE source model. The results for this case can be easily adapted to the other cases. This work extends results obtained in [SL13] for the noiseless case (WN). Two situations are considered, depending on the value of the entropy rate

$$\mathcal{H}\left(\mathsf{x}\right) = \lim_{N \to \infty} \frac{1}{N} H\left(\mathbf{x}^N\right). \tag{2.11}$$

Proposition 2.1. *Assume the presence of both additive and sensing noises and that* $\mathcal{H}\left(\mathsf{x}\right) > 0$. *Consider some arbitrary small* $\delta \in \mathbb{R}^+$. *For* $N \to \infty$, *the necessary conditions for* $P_e < \delta$ *are*

$$\mathcal{H}\left(\Theta, \mathsf{x}\right) - \mathcal{H}\left(\mathsf{x}\right) < 3\varepsilon + \delta \log Q, \tag{2.12}$$

$$H\left(p_{\mathsf{u}}\right) < \log Q, \tag{2.13}$$

and

$$\frac{M}{N} > \frac{\mathcal{H}\left(\Theta, \mathsf{x}\right) - \left(5\varepsilon + 2\delta \log Q\right)}{\log Q - H\left(p_{\mathsf{u}}\right)}, \tag{2.14}$$

where $\varepsilon \in \mathbb{R}^+$ is an arbitrary small constant.

Corollary 2.1. *Consider the same hypotheses as in Proposition 2.1 and assume now that $\mathcal{H}(\mathsf{x}) = 0$. Consider some arbitrary small $\delta \in \mathbb{R}^+$. For $N \to \infty$, the necessary condition for $P_e < \delta$ is*

$$\mathcal{H}(\Theta, \mathsf{x}) < 3\varepsilon + \delta \log Q, \tag{2.15}$$

where $\varepsilon \in \mathbb{R}^+$ is an arbitrary small constant.

In Proposition 2.1, (2.12) implies that for asymptotically exact recovery, $p\left(\mathbf{x}^N \mid \mathbf{\Theta}^N\right)$ should degenerate, almost surely, into a deterministic mapping. The condition (2.13) indicates that asymptotically exact recovery for non-deterministic sources is not possible in case of uniformly distributed additive noise. Finally, (2.14) indicates that the minimum number of required measurements depends both on the sensing and additive noises as well as on the distribution of Θ. In particular, for a given source with entropy rate $\mathcal{H}(\Theta)$, the number of necessary measurements increases with the level of the sensing noise, determined by $\mathcal{H}(\mathsf{x} \mid \Theta)$. Similarly, the number of necessary measurements increases when the additive noise gets closer to uniformly distributed. The following proof is inspired by the work in [SL13]: both additive noise and sensing noise are considered here.

Proof. From the problem setup, one has the Markov chain

$$\mathbf{\Theta}^N \leftrightarrow \mathbf{x}^N \leftrightarrow \left(\mathbf{y}^M, \mathbf{A}\right) \leftrightarrow \hat{\mathbf{\Theta}}^N, \tag{2.16}$$

from which one deduces that

$$H\left(\mathbf{\Theta}^N \mid \mathbf{x}^N\right) \leqslant H\left(\mathbf{\Theta}^N \mid \hat{\mathbf{\Theta}}^N\right), \tag{2.17}$$

and

$$H\left(\mathbf{x}^N \mid \mathbf{y}^M, \mathbf{A}\right) \leqslant H\left(\mathbf{\Theta}^N \mid \hat{\mathbf{\Theta}}^N\right). \tag{2.18}$$

Applying Fano's inequality [CT06b, Sec. 2.10], one gets

$$\begin{aligned} H\left(\mathbf{\Theta}^N \mid \hat{\mathbf{\Theta}}^N\right) &\leqslant 1 + P_{\mathrm{e}} \cdot \log\left(Q^N - 1\right) \\ &< 1 + N P_{\mathrm{e}} \log Q, \end{aligned} \tag{2.19}$$

an upper bound of P_{e} is obtained combining (2.17) and (2.19),

$$P_{\mathrm{e}} > \frac{H\left(\mathbf{\Theta}^N, \mathbf{x}^N\right) - H\left(\mathbf{x}^N\right) - 1}{N \log Q}. \tag{2.20}$$

Since $\mathbf{\Theta}^N$ and $\mathbf{x}^N$ are stationary and ergodic, for any $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\forall N > N_0$, one has

$$\begin{cases} \mathcal{H}(\Theta, \mathsf{x}) - \varepsilon < \frac{H\left(\mathbf{\Theta}^N, \mathbf{x}^N\right)}{N} < \mathcal{H}(\Theta, \mathsf{x}) + \varepsilon, \\ \mathcal{H}(\mathsf{x}) - \varepsilon < \frac{H\left(\mathbf{x}^N\right)}{N} < \mathcal{H}(\mathsf{x}) + \varepsilon, \\ \varepsilon > \frac{1}{N}. \end{cases} \tag{2.21}$$

Hence for $N > N_0$, (2.20) can be rewritten as

$$P_{\mathrm{e}} > \frac{\mathcal{H}(\Theta, \mathsf{x}) - \mathcal{H}(\mathsf{x}) - 3\varepsilon}{\log Q}. \tag{2.22}$$

For $P_{\mathrm{e}} < \delta$, one deduces (2.12) from (2.22). For δ and ε arbitrary small, (2.12) imposes $\mathcal{H}(\Theta, \mathsf{x}) = \mathcal{H}(\mathsf{x})$, meaning that Θ should be deterministic knowing x, almost surely.

From (2.18) and (2.19), one gets another lower bound for P_{e}

$$P_{\mathrm{e}} > \frac{H\left(\mathbf{x}^N \mid \mathbf{y}^M, \mathbf{A}\right) - 1}{N \log Q}. \tag{2.23}$$

The conditional entropy $H\left(\mathbf{x}^N \mid \mathbf{y}^M, \mathbf{A}\right)$ can be bounded as

$$\begin{aligned} H\left(\mathbf{x}^N \mid \mathbf{y}^M, \mathbf{A}\right) &= H\left(\mathbf{x}^N\right) - I\left(\mathbf{x}^N; \mathbf{y}^M, \mathbf{A}\right) \\ &= H\left(\mathbf{x}^N\right) - \left(I\left(\mathbf{x}^N; \mathbf{A}\right) + I\left(\mathbf{x}^N; \mathbf{y}^M \mid \mathbf{A}\right)\right) \\ &\overset{(a)}{=} H\left(\mathbf{x}^N\right) - \left(H\left(\mathbf{y}^M \mid \mathbf{A}\right) - H\left(\mathbf{y}^M \mid \mathbf{A}, \mathbf{x}^N\right)\right) \\ &\overset{(b)}{\geqslant} H\left(\mathbf{x}^N\right) - M \cdot \log Q + H\left(\mathbf{y}^M \mid \mathbf{A}, \mathbf{x}^N\right) \\ &\overset{(c)}{=} H\left(\mathbf{x}^N\right) - M \cdot \log Q + M H\left(p_{\mathsf{u}}\right), \end{aligned} \tag{2.24}$$

where (a) follows from the assumption that $\mathbf{x}^N$ and $\mathbf{A}$ are independent, (b) comes from $H\left(\mathbf{y}^M \mid \mathbf{A}\right) \leqslant H\left(\mathbf{y}^M\right) \leqslant \log \left|\mathbb{F}_Q^M\right| = M \log Q$, and (c) is because

$$H\left(\mathbf{y}^M \mid \mathbf{A}, \mathbf{x}^N\right) = H\left(\mathbf{A}\mathbf{x}^N + \mathbf{u}^M \mid \mathbf{A}, \mathbf{x}^N\right) = H\left(\mathbf{u}^M\right) = M H\left(p_{\mathsf{u}}\right). \tag{2.25}$$

Using (2.23) and (2.24), a second necessary condition for $P_{\mathrm{e}} < \delta$ is

$$\frac{H\left(\mathbf{x}^N\right) - M\left(\log Q - H\left(p_{\mathsf{u}}\right)\right) - 1}{N \log Q} < \delta. \tag{2.26}$$

For $N > N_0$, using (2.21) in (2.26) yields

$$\frac{\mathcal{H}(\mathsf{x}) - \frac{M}{N}\left(\log Q - H\left(p_{\mathsf{u}}\right)\right) - 2\varepsilon}{\log Q} < \delta. \tag{2.27}$$

Now consider two cases. In the first case, the additive noise is assumed uniformly distributed, *i.e.*,

$$H\left(p_{\mathsf{u}}\right) = \log Q; \tag{2.28}$$

the condition (2.27) becomes

$$\mathcal{H}(\mathsf{x}) < \delta \log Q + 2\varepsilon. \tag{2.29}$$

As δ can be made arbitrary small, (2.29) imposes that, for uniform additive noise, asymptotically vanishing probability of error is possible only if $\mathcal{H}(\mathsf{x})$ is arbitrary close to zero. For non-degenerate cases, *i.e.*, $\mathcal{H}(\mathsf{x}) > 0$, one obtains the necessary condition (2.13). In this second case, a lower bound

of the compression ratio M/N is obtained immediately from (2.27),

$$\frac{M}{N} > \frac{\mathcal{H}(\mathsf{x}) - (2\varepsilon + \delta \log Q)}{\log Q - H(p_{\mathsf{u}})}. \tag{2.30}$$

We can represent the condition (2.30) in terms of the joint entropy rate $\mathcal{H}(\Theta, \mathsf{x})$ by applying (2.12). Then, one gets (2.14) and Proposition 2.1 is proved.

Consider now $\mathcal{H}(\mathsf{x}) = 0$, then (2.27) holds for any value of M/N, and for any $H(p_{\mathsf{u}}) \leqslant \log Q$, since the left side of (2.27) is always negative. Hence, (2.12) is the only necessary condition for this case. Corollary 2.1 is also proved. □

With the results of the NCS noise model, one may derive the necessary conditions for the other models. If no sensing noise is considered, *i.e.*, $\mathbf{x}^N = \mathbf{\Theta}^N$, one has $H\left(\mathbf{\Theta}^N \mid \mathbf{x}^N\right) = H\left(\mathbf{x}^N \mid \mathbf{\Theta}^N\right) = 0$ and $H\left(\mathbf{\Theta}^N, \mathbf{x}^N\right) = H\left(\mathbf{\Theta}^N\right)$. If additive noise is absent, *i.e.*, $\mathbf{u}^M = \mathbf{0}$, $H\left(\mathbf{u}^M\right) = 0$. The necessary conditions for asymptotically ($N \to \infty$) vanishing probability of decoding error for each case are listed in Table 2.

TABLE 2.2: Necessary conditions for asymptotic perfect recovery in noiseless and noisy cases

Case	Necessary Condition ($\mathcal{H}(\mathsf{x}) > 0$)
WN	$\frac{M}{N} > \frac{\mathcal{H}(\Theta)}{\log Q}$, already obtained in [SL13]
NC	$\frac{M}{N} > \frac{\mathcal{H}(\Theta)}{\log Q - H(p_{\mathsf{u}})}$ and $H(p_{\mathsf{u}}) < \log Q$
NS	$\frac{M}{N} > \frac{\mathcal{H}(\Theta, \mathsf{x})}{\log Q}$ and $\mathcal{H}(\Theta \mid \mathsf{x}) = 0$
NCS	$\frac{M}{N} > \frac{\mathcal{H}(\Theta, \mathsf{x})}{\log Q - H(p_{\mathsf{u}})}$ and $H(p_{\mathsf{u}}) < \log Q$ and $\mathcal{H}(\Theta \mid \mathsf{x}) = 0$

2.4 Sufficient Condition in Absence of Sensing Noise

This section provides an upper bound of the error probability for the MAP estimation problem in absence of sensing noise (the WN and NC cases). These two cases are considered simultaneously because their proofs are similar. When the channel noise vanishes, the NC case boils down to the WN case.

2.4.1 Upper Bound of the Error Probability

Proposition 2.2. *Under MAP decoding, the asymptotic ($N \to \infty$) probability of error in absence of sensing noise can be upper bounded as*

$$P_e \leqslant P_1(\alpha) + P_2(\alpha) + 2\varepsilon, \tag{2.31}$$

where $\varepsilon \in \mathbb{R}^+$ is an arbitrarily small constant. $P_1(\alpha)$ and $P_2(\alpha)$ are defined as

$$P_1(\alpha) = 2^{-N\left(-\frac{M}{N}(H(p_{\mathsf{u}}) + \log(1-\gamma) + \varepsilon) - H_2(\alpha) - \alpha \log(Q-1) - \frac{\log(\alpha N)}{N}\right)}, \tag{2.32}$$

and

$$P_2(\alpha) = 2^{-N\left(-\mathcal{H}(\Theta)-\frac{M}{N}\left(H(p_{\mathrm{u}})+\log\left(Q^{-1}+\left(1-\frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil}\left(1-Q^{-1}\right)\right)+\varepsilon\right)-\varepsilon\right)}, \tag{2.33}$$

with $\alpha \in \mathbb{R}^+$ *and* $\alpha < 0.5$.

Proof. The proof consists of two parts. First we define the error event, and then we analyze the probability of error.

Since no sensing noise is considered, we have $\mathbf{x}^N = \mathbf{\Theta}^N$ throughout this section. The *a posteriori* pmf (2.7) becomes

$$p\left(\boldsymbol{\theta}^N \mid \mathbf{y}^M, \mathbf{A}\right) \propto \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) p\left(\mathbf{A}\right) 1_{\mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M}. \tag{2.34}$$

Suppose that $\boldsymbol{\theta}^N$ (given but unknown) is the true state vector and consider that $\mathbf{A}$ has been generated randomly. At the sink, $\mathbf{A}$ and $\mathbf{y}^M$ are known. With MAP decoding, the reconstruction $\hat{\boldsymbol{\theta}}^N$ in (2.6) is

$$\hat{\boldsymbol{\theta}}^N = \arg \max_{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N} \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) p\left(\mathbf{A}\right) 1_{\mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M}. \tag{2.35}$$

A decoding error happens if there exists a vector $\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \left\{\boldsymbol{\theta}^N\right\}$ such that

$$\sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\varphi}^N\right) p\left(\mathbf{v}^M\right) 1_{\mathbf{y}^M = \mathbf{A}\boldsymbol{\varphi}^N + \mathbf{v}^M} \geqslant \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) 1_{\mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M}. \tag{2.36}$$

For fixed $\mathbf{y}^M$, $\mathbf{A}$, and $\boldsymbol{\theta}^N$, there is exactly one vector $\mathbf{u}^M$ such that $\mathbf{u}^M = \mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N$. Hence the right side of (2.36) can be represented as $p_{\mathbf{\Theta}^N}\left(\boldsymbol{\theta}^N\right) p_{\mathbf{u}^M}\left(\mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N\right)$. The subscripts for the pmfs are introduced to avoid any ambiguity of notations. Then (2.36) is equivalent to

$$p_{\mathbf{\Theta}^N}\left(\boldsymbol{\varphi}^N\right) p_{\mathbf{u}^M}\left(\mathbf{y}^M - \mathbf{A}\boldsymbol{\varphi}^N\right) \geqslant p_{\mathbf{\Theta}^N}\left(\boldsymbol{\theta}^N\right) p_{\mathbf{u}^M}\left(\mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N\right). \tag{2.37}$$

An alternative way to state the error event can be: For a given realization $\mathbf{\Theta}^N = \boldsymbol{\theta}^N$, which implies the realization $\mathbf{u}^M = \mathbf{u}^M = \mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N$, there exists a pair $(\boldsymbol{\varphi}^N, \mathbf{v}^M) \in \mathbb{F}_Q^N \times \mathbb{F}_Q^M$ such that

$$\begin{cases} \boldsymbol{\varphi}^N \neq \boldsymbol{\theta}^N, \\ \mathbf{A}\boldsymbol{\varphi}^N + \mathbf{v}^M = \mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M, \\ p\left(\boldsymbol{\varphi}^N\right) p\left(\mathbf{v}^M\right) \geqslant p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right). \end{cases} \tag{2.38}$$

From conditions (2.38), one concludes that the MAP decoder is equivalent to the maximum Q-probability decoder [Csi82] in the NC case.

An upper bound of the error probability is now derived. For a fixed $\boldsymbol{\theta}^N$ and $\mathbf{u}^M$, the conditional error probability is denoted by $\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\}$. The average error probability is

$$P_{\mathrm{e}} = \sum_{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N} \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p\left(\boldsymbol{\theta}^N, \mathbf{u}^M\right) \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\}. \tag{2.39}$$

Weak typicality is instrumental in the following proofs. The notations of [Yeu04, Definition 4.2] are extended to stationary and ergodic sources. For any positive real number ε and some integer $N > 0$, the weakly typical set $\mathcal{A}^N_{[\Theta]_\varepsilon} \subset \mathbb{F}_Q^N$ for a stationary and ergodic source $\mathbf{\Theta}^N$ is the set of vectors $\boldsymbol{\theta}^N \in \mathbb{F}_Q^N$ satisfying

$$\left| -\frac{1}{N} \log p\left(\boldsymbol{\theta}^N\right) - \mathcal{H}\left(\Theta\right) \right| \leqslant \varepsilon, \tag{2.40}$$

where $\mathcal{H}\left(\Theta\right)$ is the entropy rate of the source. Similarly, for the noise vector $\mathbf{u}^M$, define

$$\mathcal{A}^M_{[\mathsf{u}]_\varepsilon} = \left\{ \mathbf{u}^M \in \mathbb{F}_Q^M : \left| -\frac{1}{M} \log p\left(\mathbf{u}^M\right) - H\left(p_{\mathsf{u}}\right) \right| \leqslant \varepsilon \right\}. \tag{2.41}$$

Recall that the entries of $\mathbf{u}^M$ are uncorrelated, so $\mathcal{H}\left(\mathsf{u}\right) = H\left(p_{\mathsf{u}}\right)$. Thanks to Shannon-McMillan-Breiman theorem [CT06b, Sec. 16.8], the pmf of the general stationary and ergodic source converges. In other words, for any $\varepsilon > 0$, there exists N_ε and M_ε such that for all $N > N_\varepsilon$ and $M > M_\varepsilon$,

$$\Pr\left\{ \left| -\frac{1}{N} \log p\left(\mathbf{\Theta}^N\right) - \mathcal{H}\left(\Theta\right) \right| \leqslant \varepsilon \right\} \geqslant 1 - \varepsilon, \tag{2.42}$$

and

$$\Pr\left\{ \left| -\frac{1}{M} \log p\left(\mathbf{u}^M\right) - \mathcal{H}\left(p_{\mathsf{u}}\right) \right| \leqslant \varepsilon \right\} \geqslant 1 - \varepsilon. \tag{2.43}$$

We can make ε arbitrary close to zero as $N \to \infty$ and $M \to \infty$. A sandwich proof of this theorem is proposed in [CT06b, Sec. 16.8]. For the sparse and uncorrelated source defined in (2.1), $\mathcal{H}\left(\Theta\right)$ is equal to $H\left(p_\Theta\right)$, the entropy of a single source. The entropy rate of the StM source is the conditional entropy $H\left(\Theta_{n+r} \mid \mathbf{\Theta}_n^{n+r-1}\right)$.

From (2.42) and (2.43), one has $\Pr\left\{\mathbf{\Theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}\right\} \geqslant 1 - \varepsilon$ and $\Pr\left\{\mathbf{u}^M \in \mathcal{A}^M_{[\mathsf{u}]_\varepsilon}\right\} \geqslant 1 - \varepsilon$ for $N > N_\varepsilon$ and $M > M_\varepsilon$. It implies that, for N and M sufficiently large, $\mathbf{\Theta}^N$ and $\mathbf{u}^M$ belong to the weakly typical set $\mathcal{A}^N_{[\Theta]_\varepsilon}$ and $\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}$, almost surely. With respect to the typicality, $\mathbb{F}_Q^N \times \mathbb{F}_Q^M$ can be divided into two parts. Define the sets $\mathcal{U}$ and $\mathcal{U}^\mathrm{c}$ for the pair of vectors $\left(\boldsymbol{\theta}^N, \mathbf{u}^M\right)$, such that $\mathcal{U} \cup \mathcal{U}^\mathrm{c} = \mathbb{F}_Q^N \times \mathbb{F}_Q^M$ and

$$\mathcal{U} = \left\{ \boldsymbol{\theta}^N \in \mathbb{F}_Q^N,\, \mathbf{u}^M \in \mathbb{F}_Q^M : \boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \text{ and } \mathbf{u}^M \in \mathcal{A}^M_{[\mathsf{u}]_\varepsilon} \right\}, \tag{2.44}$$

$$\mathcal{U}^\mathrm{c} = \left\{ \boldsymbol{\theta}^N \in \mathbb{F}_Q^N,\, \mathbf{u}^M \in \mathbb{F}_Q^M : \boldsymbol{\theta}^N \notin \mathcal{A}^N_{[\Theta]_\varepsilon} \text{ or } \mathbf{u}^M \notin \mathcal{A}^M_{[\mathsf{u}]_\varepsilon} \right\}. \tag{2.45}$$

$\mathcal{U}$ is the joint typical set for $(\boldsymbol{\theta}^N, \mathbf{u}^M)$, due to the independence of $\boldsymbol{\Theta}^N$ and $\mathbf{u}^M$. The error probability can be bounded as

$$\begin{aligned} P_{\mathrm{e}} &= \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \cdot \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\} \\ &\quad + \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}^{\mathrm{c}}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \cdot \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\} \\ &\overset{(a)}{\leqslant} \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \cdot \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\} + \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}^{\mathrm{c}}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \\ &\overset{(b)}{\leqslant} \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \cdot \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\} + 2\varepsilon, \end{aligned} \tag{2.46}$$

where (a) comes from $\Pr\left(\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right) \leqslant 1$ and (b) follows from the fact that

$$\begin{aligned} \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}^{\mathrm{c}}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) &= 1 - \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \\ &= 1 - \sum_{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]\varepsilon}} p\left(\boldsymbol{\theta}^N\right) \sum_{\mathbf{u}^M \in \mathcal{A}^M_{[\mathrm{u}]\varepsilon}} p\left(\mathbf{u}^M\right) \\ &\leqslant 1 - (1-\varepsilon)(1-\varepsilon) \leqslant 2\varepsilon. \end{aligned} \tag{2.47}$$

Since $\mathbf{A}$ is generated randomly, define the random event

$$\mathcal{E}\left(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M\right) = \left\{\mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M = \mathbf{A}\boldsymbol{\varphi}^N + \mathbf{v}^M\right\}, \tag{2.48}$$

where $\left(\boldsymbol{\theta}^N, \mathbf{u}^M\right)$ is the realization of the environment state, and $(\boldsymbol{\varphi}^N, \mathbf{v}^M)$ is the potential reconstruction result. Conditioned on $\left(\boldsymbol{\theta}^N, \mathbf{u}^M\right)$, $\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\}$ is in fact the probability of the union of the events $\mathcal{E}\left(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M\right)$ with all the parameter pairs $(\boldsymbol{\varphi}^N, \mathbf{v}^M) \in \mathbb{F}_Q^N \times \mathbb{F}_Q^M$ such that $\boldsymbol{\varphi}^N \neq \boldsymbol{\theta}^N$ and $p\left(\boldsymbol{\varphi}^N\right) p\left(\mathbf{v}^M\right) \geqslant p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right)$, see (2.38). The conditional error probability can then be rewritten as

$$\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{u}^M\right\} = \Pr\left\{ \bigcup_{\substack{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\},\, \mathbf{v}^M \in \mathbb{F}_Q^M : \\ p(\boldsymbol{\varphi}^N) p(\mathbf{v}^M) \geqslant p(\boldsymbol{\theta}^N) p(\mathbf{u}^M)}} \mathcal{E}\left(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M\right) \right\}. \tag{2.49}$$

Introducing (2.49) in (2.46) and applying the union bound yields

$$\begin{aligned} P_{\mathrm{e}} &\leqslant \sum_{\left(\boldsymbol{\theta}^N,\mathbf{u}^M\right)\in\mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \sum_{\substack{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\},\, \mathbf{v}^M\in\mathbb{F}_Q^M:\\ p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M)\geqslant p(\boldsymbol{\theta}^N)p(\mathbf{u}^M)}} \Pr\left\{\mathcal{E}\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right)\right\}+2\varepsilon \\ &= \sum_{\left(\boldsymbol{\theta}^N,\mathbf{u}^M\right)\in\mathcal{U}} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{u}^M\right) \\ &\qquad \cdot \sum_{\substack{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}\\ \mathbf{v}^M\in\mathbb{F}_Q^M}} \Phi\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right) \Pr\left\{\mathcal{E}\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right)\right\}+2\varepsilon, \end{aligned} \tag{2.50}$$

where

$$\Phi\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right)=\begin{cases}1 & \text{if } p\left(\boldsymbol{\varphi}^N\right)p\left(\mathbf{v}^M\right)\geqslant p\left(\boldsymbol{\theta}^N\right)p\left(\mathbf{u}^M\right),\\ 0 & \text{if } p\left(\boldsymbol{\varphi}^N\right)p\left(\mathbf{v}^M\right)< p\left(\boldsymbol{\theta}^N\right)p\left(\mathbf{u}^M\right).\end{cases} \tag{2.51}$$

Now consider the following lemma.

Lemma 2.1. *Consider some $s\in\mathbb{R}^+$ with $s\leqslant 1$. For any $\boldsymbol{\theta}^N,\boldsymbol{\varphi}^N$ in $\mathbb{F}_Q^N$ and $\mathbf{u}^M,\mathbf{v}^M$ in $\mathbb{F}_Q^M$, the following inequality holds,*

$$\Phi\left(\boldsymbol{\theta}^N,\boldsymbol{\varphi}^N,\mathbf{u}^M,\mathbf{v}^M\right)\leqslant\left(\frac{p\left(\boldsymbol{\varphi}^N\right)p\left(\mathbf{v}^M\right)}{p\left(\boldsymbol{\theta}^N\right)p\left(\mathbf{u}^M\right)}\right)^s. \tag{2.52}$$

Lemma 2.1 is a part of Gallager's derivation of error exponents in [Gal68, Sec. 5.6]. Introducing (2.52) with $s=1$ into (2.50), one gets

$$P_{\mathrm{e}}\leqslant \sum_{\left(\boldsymbol{\theta}^N,\mathbf{u}^M\right)\in\mathcal{U}} \sum_{\substack{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}\\ \mathbf{v}^M\in\mathbb{F}_Q^M}} p\left(\boldsymbol{\varphi}^N\right)p\left(\mathbf{v}^M\right)\Pr\left\{\mathcal{E}\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right)\right\}+2\varepsilon. \tag{2.53}$$

In (2.53),

$$\Pr\left\{\mathcal{E}\left(\boldsymbol{\theta}^N,\mathbf{u}^M;\boldsymbol{\varphi}^N,\mathbf{v}^M\right)\right\}=\Pr\left\{\mathbf{A}\boldsymbol{\mu}^N=\mathbf{s}^M\mid\boldsymbol{\mu}^N\neq\mathbf{0},\mathbf{s}^M\right\} \tag{2.54}$$

with $\boldsymbol{\mu}^N=\boldsymbol{\varphi}^N-\boldsymbol{\theta}^N\in\mathbb{F}_Q^N\setminus\{\mathbf{0}\}$, and $\mathbf{s}^M=\mathbf{u}^M-\mathbf{v}^M\in\mathbb{F}_Q^M$. This probability depends on the sparsity of $\boldsymbol{\mu}^N$ and of $\mathbf{s}^M$, let $d_1=\left\|\boldsymbol{\mu}^N\right\|_0$ and $d_2=\left\|\mathbf{s}^M\right\|_0$. Both d_1 and d_2 are integers such that $1\leqslant d_1\leqslant N$ and $0\leqslant d_2\leqslant M$. Define the multivariable function

$$f\left(d_1,d_2\right)=\Pr\left\{\mathbf{A}\boldsymbol{\mu}^N=\mathbf{s}^M\mid\left\|\boldsymbol{\mu}^N\right\|_0=d_1,\left\|\mathbf{s}^M\right\|_0=d_2\right\}, \tag{2.55}$$

which depends on γ, Q, and M of the random matrix $\mathbf{A}$.

Lemma 2.2. *The function $f\left(d_1,d_2\right)$, defined in (2.55), is non-increasing in d_2 for a given d_1 and*

$$f\left(d_1,d_2\right)\leqslant f\left(d_1,0\right)\leqslant\left(Q^{-1}+\left(1-\frac{\gamma}{1-Q^{-1}}\right)^{d_1}\left(1-Q^{-1}\right)\right)^M. \tag{2.56}$$

Moreover $f(d_1, 0)$ *is non-increasing in* d_1 *and*

$$f(d_1, 0) \leqslant f(1, 0) \leqslant (1-\gamma)^M . \tag{2.57}$$

If $\gamma_1 = \cdots = \gamma_N = 1 - Q^{-1}$*, which corresponds to a uniformly distributed network coding matrix,*

$$f(d_1, d_2) = Q^{-M} \tag{2.58}$$

is constant.

See Section A.1 for the proof details. Using Lemma 2.2, (2.53) can be expressed as

$$\begin{aligned}
P_{\mathrm{e}} &\overset{(a)}{\leqslant} \sum_{d_1=1}^{N} \sum_{d_2=0}^{M} \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N : \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1 \\ \mathbf{v}^M \in \mathbb{F}_Q^M : \|\mathbf{u}^M - \mathbf{v}^M\|_0 = d_2}} p\left(\boldsymbol{\varphi}^N\right) p\left(\mathbf{v}^M\right) f(d_1, d_2; \gamma, Q, M) + 2\varepsilon \\
&\overset{(b)}{\leqslant} \sum_{d_1=1}^{N} \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N : \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p\left(\boldsymbol{\varphi}^N\right) f(d_1, 0; \gamma, Q, M) \left(\sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p\left(\mathbf{v}^M\right) \right) + 2\varepsilon \\
&\overset{(c)}{\leqslant} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N : \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p\left(\boldsymbol{\varphi}^N\right) f(1, 0; \gamma, Q, M) \\
&\qquad + \sum_{d_1=\lceil \alpha N \rceil}^{N} \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N : \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p\left(\boldsymbol{\varphi}^N\right) f(\lceil \alpha N \rceil, 0; \gamma, Q, M) + 2\varepsilon,
\end{aligned} \tag{2.59}$$

where (a) is by the classification of $\boldsymbol{\varphi}^N$ and $\mathbf{v}^M$ according to the ℓ_0 norm of their difference with $\boldsymbol{\theta}^N$ and $\mathbf{u}^M$ respectively and (b) is obtained using the bound (2.56) and using $\sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p\left(\mathbf{v}^M\right) = 1$. The splitting in (c) permits $f(d_1, 0; \gamma, Q, M)$ to be bounded in different cases; this idea comes from [TBD12] and is also meaningful here. The parameter α is a positive real number with $0 < \alpha < 0.5$. The best choice of α is discussed in Section 2.4.2. The two terms in (2.59), denoted by $P_{\mathcal{U}_1}(\alpha)$ and

$P_{\mathcal{U}_2}(\alpha)$, need to be considered separately. For the first term $P_{\mathcal{U}_1}(\alpha)$, we have

$$\begin{aligned}
P_{\mathcal{U}_1}(\alpha) &= f(1,0;\gamma,Q,M) \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\substack{(\boldsymbol{\theta}^N,\mathbf{u}^M)\in\mathcal{U} \\ \boldsymbol{\varphi}^N\in\mathbb{F}_Q^N:\|\boldsymbol{\varphi}^N-\boldsymbol{\theta}^N\|_0=d_1}} p\left(\boldsymbol{\varphi}^N\right) \\
&\overset{(a)}{=} (1-\gamma)^M \sum_{\mathbf{u}^M\in\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N\right) \sum_{\substack{\boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}: \\ \|\boldsymbol{\theta}^N-\boldsymbol{\varphi}^N\|_0=d_1}} 1 \\
&\overset{(b)}{\leqslant} (1-\gamma)^M \sum_{\mathbf{u}^M\in\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N\right) \left|\left\{\boldsymbol{\theta}^N\in\mathbb{F}_Q^N : \left\|\boldsymbol{\theta}^N-\boldsymbol{\varphi}^N\right\|_0=d_1\right\}\right| \\
&\overset{(c)}{\leqslant} (1-\gamma)^M \sum_{\mathbf{u}^M\in\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} 2^{NH_2\left(\frac{d_1}{N}\right)} (Q-1)^{d_1} \\
&\overset{(d)}{\leqslant} (1-\gamma)^M \cdot \left|\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}\right| \cdot \alpha N \cdot 2^{NH_2(\alpha)} (Q-1)^{\alpha N} \\
&\overset{(e)}{\leqslant} 2^{-N\left(-\frac{M}{N}(H(p_\mathsf{u})+\log(1-\gamma)+\varepsilon)-H_2(\alpha)-\alpha\log(Q-1)-\frac{\log(\alpha N)}{N}\right)} = P_1(\alpha)
\end{aligned} \tag{2.60}$$

where (a) is by changing the order of summation and (b) is obtained considering all $\boldsymbol{\theta}^N\in\mathbb{F}_Q^N$ and not only typical sequences. The bound (c) is obtained noticing that

$$\begin{aligned}
\left|\left\{\boldsymbol{\theta}^N\in\mathbb{F}_Q^N : \left\|\boldsymbol{\theta}^N-\boldsymbol{\varphi}^N\right\|_0=d_1\right\}\right| &= \binom{N}{d_1}(Q-1)^{d_1} \\
&\leqslant 2^{NH_2\left(\frac{d_1}{N}\right)}(Q-1)^{d_1},
\end{aligned} \tag{2.61}$$

where $H_2(p)$ denotes the entropy of a Bernoulli-p source and $\sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N\right)=1$; (d) is because of the monotonicity of the function $H_2\left(\frac{d_1}{N}\right)$, which is increasing in d_1 as $d_1\leqslant\lfloor\alpha N\rfloor<N/2$; (e) comes from [CT06b, Theorem 3.1.2], the upper bound of the size of $\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}$, *i.e.*,

$$\left|\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}\right| \leqslant 2^{M(H(p_\mathsf{u})+\varepsilon)}, \tag{2.62}$$

for $M>M_\varepsilon$. Similarly, for $N>N_\varepsilon$, one has

$$\left|\mathcal{A}^N_{[\Theta]_\varepsilon}\right| \leqslant 2^{N(\mathcal{H}(\Theta)+\varepsilon)}. \tag{2.63}$$

Now we turn to $P_{\mathcal{U}_2}(\alpha)$,

$$\begin{aligned} P_{\mathcal{U}_2}(\alpha) &= \sum_{d_1=\lceil \alpha N\rceil}^{N} \sum_{\substack{(\boldsymbol{\theta}^N,\mathbf{u}^M)\in\mathcal{U} \\ \boldsymbol{\varphi}^N\in\mathbb{F}_Q^N:\|\boldsymbol{\varphi}^N-\boldsymbol{\theta}^N\|_0=d_1}} p\left(\boldsymbol{\varphi}^N\right) f\left(\lceil \alpha N\rceil, 0;\gamma,Q,M\right) \\ &\overset{(a)}{\leqslant} \sum_{(\boldsymbol{\theta}^N,\mathbf{u}^M)\in\mathcal{U}} \sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N\right) f\left(\lceil \alpha N\rceil, 0;\gamma,Q,M\right) \\ &= \left|\mathcal{A}^N_{[\Theta]_\varepsilon}\right| \cdot \left|\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}\right| \cdot \left(Q^{-1}+\left(1-\frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil}\left(1-Q^{-1}\right)\right)^M \\ &\overset{(b)}{\leqslant} 2^{-N\left(-\mathcal{H}(\Theta)-\frac{M}{N}\left(H(p_{\mathsf{u}})+\log\left(Q^{-1}+\left(1-\frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil}\left(1-Q^{-1}\right)\right)+\varepsilon\right)-\varepsilon\right)} = P_2(\alpha), \end{aligned} \tag{2.64}$$

where (a) is by ignoring the constraint that $\left\|\boldsymbol{\varphi}^N-\boldsymbol{\theta}^N\right\|_0 = d_1$, and (b) is by the upper bounds of $\left|\mathcal{A}^N_{[\Theta]_\varepsilon}\right|$ and $\left|\mathcal{A}^M_{[\mathsf{u}]_\varepsilon}\right|$, as before. Equations (2.59), (2.60), and (2.64) complete the proof. □

2.4.2 Sufficient Condition

In this section, sufficient conditions for the WN and NC cases are derived to get a vanishing upper bound to the error probability.

Proposition 2.3. *Assume the absence of sensing noise and consider a network coding matrix with sparsity factor γ. For some $\delta\in\mathbb{R}^+$ (which may be taken arbitrary close to zero), there exists small positive real numbers ε, ξ, and integers N_δ, M_ε such that $\forall N > N_\delta$ and $M > M_\varepsilon$, if the following conditions hold*

- *the additive noise is not uniformly distributed, i.e.,*

$$H(p_{\mathsf{u}}) < \log Q - \xi, \tag{2.65}$$

- *the sparsity factor is lower bounded*

$$\gamma > 1 - 2^{-H(p_{\mathsf{u}})-\varepsilon}, \tag{2.66}$$

- *the compression ratio M/N satisfies*

$$\frac{M}{N} > \frac{\mathcal{H}(\Theta)+\varepsilon}{\log Q - H(p_{\mathsf{u}}) - \xi}. \tag{2.67}$$

Then, one has $P_e \leqslant \delta$ using MAP decoding. As $N\to\infty$ and $M\to\infty$, ε and ξ can be chosen arbitrary close to zero.

Proof. Both $P_1(\alpha)$ and $P_2(\alpha)$ need to be vanishing for increasing N and M. The exponent of each term is considered respectively. Define, from (2.60),

$$E_1^{\mathrm{NC}} = -\frac{M}{N}\left(H(p_{\mathsf{u}}) + \log(1-\gamma) + \varepsilon\right) - H_2(\alpha) - \alpha\log(Q-1) - \frac{\log(\alpha N)}{N}. \tag{2.68}$$

Then $\lim_{N\to\infty} 2^{-NE_1^{\mathrm{NC}}} = 0$ if $E_1^{\mathrm{NC}} > 0$. Thus, if $E_1^{\mathrm{NC}} > 0$, for any $\tau_1 \in \mathbb{R}^+$ arbitrarily small, $\exists N_{\tau_1}$ such that $\forall N > N_{\tau_1}$, one has $P_1(\alpha) < \tau_1$.

Notice that if $H(p_{\mathsf{u}}) + \log(1-\gamma) + \varepsilon \geqslant 0$, E_1^{NC} is negative, thus one should first have

$$H(p_{\mathsf{u}}) + \log(1-\gamma) + \varepsilon < 0, \tag{2.69}$$

leading to (2.66). With this condition, $E_1^{\mathrm{NC}} > 0$ leads to

$$\frac{M}{N} > \frac{H_2(\alpha) + \alpha\log(Q-1) + \frac{\log(\alpha N)}{N}}{\log\frac{1}{1-\gamma} - H(p_{\mathsf{u}}) - \varepsilon}. \tag{2.70}$$

Similarly, define from (2.64)

$$E_2^{\mathrm{NC}} = -\mathcal{H}(\Theta) - \frac{M}{N}\left(H(p_{\mathsf{u}}) + \log\left(Q^{-1} + \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil}\left(1-Q^{-1}\right)\right) + \varepsilon\right) - \varepsilon. \tag{2.71}$$

Again, if $E_2^{\mathrm{NC}} > 0$, for any $\tau_2 \in \mathbb{R}^+$ arbitrarily small, $\exists N_{\tau_2} \in \mathbb{N}^+$ such that $\forall N > N_{\tau_2}$, one has $P_2(\alpha) < \tau_2$. Since $0 < \gamma \leqslant 1 - Q^{-1}$, one gets $0 \leqslant 1 - \frac{\gamma}{1-Q^{-1}} < 1$ and

$$\lim_{N\to\infty}\left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil} = 0. \tag{2.72}$$

Thus for $\sigma \in \mathbb{R}^+$ arbitrarily small, there exists an N_σ such that for $\forall N > N_\sigma$,

$$\left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N\rceil}\left(1-Q^{-1}\right) < \sigma Q^{-1}. \tag{2.73}$$

Hence E_2^{NC} in (2.71) can be lower bounded by

$$E_2^{\mathrm{NC}} > -\mathcal{H}(\Theta) - \frac{M}{N}\left(H(p_{\mathsf{u}}) + \log\left(Q^{-1} + \sigma Q^{-1}\right) + \varepsilon\right) - \varepsilon, \tag{2.74}$$

for $N > N_\sigma$. If this lower bound is positive, then E_2^{NC} is positive. Again, if $H(p_{\mathsf{u}}) - \log Q + \log(1+\sigma) + \varepsilon \leqslant 0$, one obtains a negative lower bound for E_2^{NC} from (2.74). Thus, one deduces (2.65) in Proposition 2.3, with

$$\xi = \log(1+\sigma) + \varepsilon. \tag{2.75}$$

From (2.65), to get a positive lower bound for (2.74), one should have

$$\frac{M}{N} > \frac{\mathcal{H}(\Theta) + \varepsilon}{\log Q - H(p_{\mathsf{u}}) - \log(1+\sigma) - \varepsilon}. \tag{2.76}$$

From (2.76) and (2.75), with $\xi \to 0$ as $N \to \infty$, one gets (2.67) in Proposition 2.3.

From (2.70) and (2.76), one obtains

$$\frac{M}{N} > \max\left\{\frac{H_2(\alpha) + \alpha\log(Q-1) + \frac{\log(\alpha N)}{N}}{\log\frac{1}{1-\gamma} - H(p_{\mathsf{u}}) - \varepsilon}, \frac{\mathcal{H}(\Theta) + \varepsilon}{\log Q - H(p_{\mathsf{u}}) - \xi}\right\}. \tag{2.77}$$

The value of α should be chosen such that the lower bound (2.77) on M/N is minimum. One may compare (2.77) with the necessary condition (2.14). The second term of (2.77) is similar to (2.14), since both ξ and ε can be made arbitrarily close to 0 as $N \to \infty$. The best value for α has thus to be such that

$$\frac{H_2(\alpha) + \alpha \log(Q-1) + \frac{\log(\alpha N)}{N}}{\log \frac{1}{1-\gamma} - H(p_{\mathrm{u}}) - \varepsilon} \leqslant \frac{\mathcal{H}(\Theta) + \varepsilon}{\log Q - H(p_{\mathrm{u}}) - \xi}. \tag{2.78}$$

The function $H_2(\alpha) + \alpha \log(Q-1)$ is increasing when $\alpha \in]0, 0.5[$ and tends to 0 as $\alpha \to 0$. The term $\log(\alpha N)/N$ is also negligible for N large. Thus, there always exists some α satisfying (2.78). Since the speed of convergence of ξ is affected by α, we choose the largest α that satisfies (2.78). Finally, the sufficient condition (2.67) is obtained for M/N.

From (2.31), one may conclude that

$$P_{\mathrm{e}} \leqslant \tau_1 + \tau_2 + 2\varepsilon. \tag{2.79}$$

To ensure $P_{\mathrm{e}} < \delta$, we should choose τ_1, τ_2, and ε to satisfy $\tau_1 + \tau_2 + 2\varepsilon < \delta$. Then a proper value of σ, which depends on τ_2 and ε, can be chosen. At last, ξ is obtained from (2.75). With these well determined parameters, if all the three conditions in Proposition 2.3 hold, there exists integers N_ε, N_{τ_1}, N_{τ_2}, and N_σ, such that for any

$$N > N_\delta = \max\{N_\varepsilon, N_{\tau_1}, N_{\tau_2} N_\sigma\}, \tag{2.80}$$

and $M > M_\varepsilon$, one has $P_{\mathrm{e}} < \delta$. □

2.4.3 Discussion and Numerical Results

In [DM09, Eq. (24)], considering a sparse and iid source, a uniformly distributed random matrix $\mathbf{A}$, and the minimum empirical entropy decoder, the following error exponent in the case NC is obtained

$$E_0^{\mathrm{NC}} = \min_{p,q} D(p \parallel p_\Theta) + \frac{M}{N} D(q \parallel p_{\mathrm{u}}) + \left| \frac{M}{N} \log Q - H(p) - \frac{M}{N} H(q) \right|^+, \tag{2.81}$$

where $D(\cdot \parallel \cdot)$ denotes the relative entropy between two distributions and $|\cdot|^+ = \max\{0, \cdot\}$. In parallel, [Ho+06] proposed an approach to prove that the upper bound for the probability of decoding error P_{e} under minimum empirical entropy decoding is equal to that of the maximum Q-probability decoder. As discussed in Section 2.4.1, in the WN and NC cases, the MAP decoder in the considered context is equivalent to the maximum Q-probability decoder. As a consequence, (2.81) is also the error exponent of the MAP decoder. A proof for (2.81) using the method of types needs to make some assumptions on the topology of the considered sensor network to specify the type of $\boldsymbol{\theta}^N$. For correlated sources, one can extend (2.81) considering a Markov model, and use higher-order types, leading to cumbersome derivations.

From (2.81), provided that $E_0^{\mathrm{NC}} > 0$, P_{e} tends to 0 as N increases. E_0^{NC} cannot be negative and $E_0^{\mathrm{NC}} = 0$ if and only if

$$\begin{cases} D\left(p \parallel p_{\Theta}\right) = 0, \\ D\left(q \parallel p_{\mathsf{u}}\right) = 0, \\ \frac{M}{N} \log Q - H\left(p\right) - \frac{M}{N} H\left(q\right) \leqslant 0. \end{cases} \tag{2.82}$$

Thus, (2.82) implies that $\frac{M}{N} \log Q - H\left(p_{\Theta}\right) - \frac{M}{N} H\left(p_{\mathsf{u}}\right) \leqslant 0$. Then, a necessary and sufficient condition to have $E_0^{\mathrm{NC}} > 0$ is $\frac{M}{N} \log Q - H\left(p_{\Theta}\right) - \frac{M}{N} H\left(p_{\mathsf{u}}\right) > 0$, which is the same as (2.67) with $\gamma = 1 - Q^{-1}$ (corresponding to $\mathbf{A}$ uniformly distributed). The proof using weak typicality leads to the same results (in terms of sufficient condition for having asymptotically vanishing P_{e}) as the technique in [DM09].

In the noiseless case, since γ can be chosen arbitrarily small, the necessary condition in Proposition 2.1 and the sufficient condition in Proposition 2.3 asymptotically coincide. This confirms the numerical results obtained in [SL13]. In the NC case, the difference between the two conditions comes from the constraint linking γ and the entropy of the additive noise. In Section 2.3, the structure of $\mathbf{A}$ was not considered and no condition on γ has been obtained. The lower bound on γ implies that $\mathbf{A}$ should be dense enough to fight against the noise. Since the additive noise is iid, for a given probability of having one entry of $\mathbf{u}^M$ non-zero, *i.e.*, $\Pr\left(\mathsf{u} \neq 0\right)$, the entropy $H\left(p_{\mathsf{u}}\right)$ is maximized when $p_{\mathsf{u}}\left(q\right) = \Pr\left(\mathsf{u} \neq 0\right) / \left(Q - 1\right)$ for any $q \in \mathbb{F}_Q \setminus \{0\}$. This corresponds to the worst noise in terms of compression efficiency.

Figure 2.3 represents the lower bound of γ as a function of $\Pr\left(\mathsf{u} \neq 0\right)$, ranging from 10^{-5} to 10^{-1}, for different values of Q. There is almost no requirement on γ when $\Pr\left(\mathsf{u} \neq 0\right) \leqslant 5 \times 10^{-4}$. For a given noise level, a larger size of the finite field needs a denser network coding matrix. Figure 2.4 shows the influence of the additive noise on the optimum compression ratio. The lower bound of M/N is represented as a function of $\mathcal{H}\left(\Theta\right) / \log Q$, for different values of Q and for different values of $\Pr\left(\mathsf{u} \neq 0\right)$.

2.5 Sufficient Condition in Presence of Sensing Noise

This section performs an achievability study in presence of sensing noise by considering the conditional pmf $p_{\mathsf{x}|\Theta}$. The additive noise $\mathbf{u}^M$ is first neglected to simplify the problem (NS case). The extension to the NCS case is easily obtained from the NS case. Assume that $\boldsymbol{\theta}^N$ is the true state vector and that $\mathbf{x}^N$ represents the measurements of the sensors. The sink receives $\mathbf{y}^M = \mathbf{A}\mathbf{x}^N$. The *a posteriori* pmf (2.7) can be written as

$$p\left(\boldsymbol{\theta}^N \mid \mathbf{y}^M, \mathbf{A}\right) \propto \sum_{\mathbf{z}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N\right) p\left(\mathbf{z}^N \mid \boldsymbol{\theta}^N\right) 1_{\mathbf{y}^M = \mathbf{A}\mathbf{z}^N}. \tag{2.83}$$

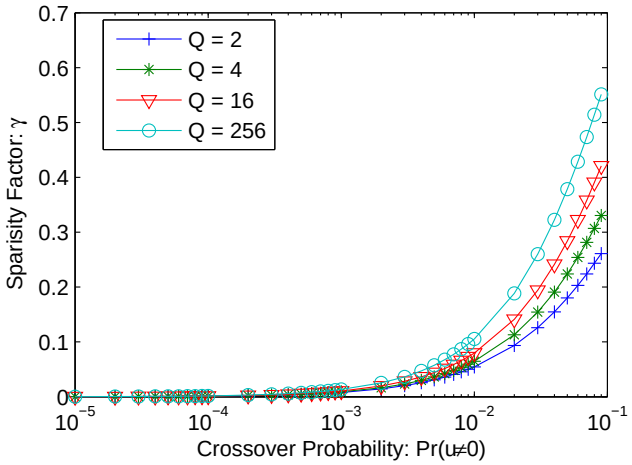

FIGURE 2.3: Lower bound of γ to achieve the optimum compression ratio for $N \to \infty$, according to (2.66)

In the case of MAP estimation, an error occurs if there exists a vector $\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}$ such that

$$\sum_{\mathbf{z}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N, \mathbf{z}^N\right) 1_{\mathbf{y}^M = \mathbf{A}\mathbf{z}^N} \leqslant \sum_{\mathbf{z}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N, \mathbf{z}^N\right) 1_{\mathbf{y}^M = \mathbf{A}\mathbf{z}^N}. \tag{2.84}$$

$\boldsymbol{\theta}^N$ and $\mathbf{x}^N$ are considered as fixed, but unknown. The decoder has knowledge of $\mathbf{A}$ and $\mathbf{y}^M = \mathbf{A}\mathbf{x}^N$, thus an alternative way to express (2.84) is

$$\sum_{\mathbf{z}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N, \mathbf{z}^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}^N} \leqslant \sum_{\mathbf{z}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N, \mathbf{z}^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}^N}. \tag{2.85}$$

2.5.1 Achievability Study

We begin with the extension of the basic weakly typical set as introduced in Section 2.4.1. For any $\varepsilon > 0$ and $N \in \mathbb{N}^+$, based on $\mathcal{A}^N_{[\Theta]_\varepsilon}$ for $\boldsymbol{\theta}^N$, one defines the weakly conditional typical set $\mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right)$ for $\mathbf{x}^N$, which is conditionally distributed with respect to $p_{\mathsf{x}|\Theta}$, with $\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}$,

$$\mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right) = \left\{\mathbf{x}^N \in \mathbb{F}_Q^N \text{ such that } \left|-\frac{1}{N}\log p\left(\mathbf{x}^N \mid \boldsymbol{\theta}^N\right) - \mathcal{H}\left(\mathsf{x} \mid \Theta\right)\right| \leqslant \varepsilon\right\}. \tag{2.86}$$

Since $\mathcal{H}\left(\Theta, \mathsf{x}\right) = \mathcal{H}\left(\Theta\right) + \mathcal{H}\left(\mathsf{x} \mid \Theta\right)$, if $\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}$ and $\mathbf{x}^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right)$, then $\left(\boldsymbol{\theta}^N, \mathbf{x}^N\right) \in \mathcal{A}^N_{[\Theta,\mathsf{x}]2\varepsilon}$ by consistency, where $\mathcal{A}^N_{[\Theta,\mathsf{x}]2\varepsilon}$ denotes the weakly joint typical set, *i.e.*, the set of pairs $\left(\boldsymbol{\theta}^N, \mathbf{x}^N\right) \in \mathbb{F}_Q^N \times \mathbb{F}_Q^N$ such that

$$\left|-\frac{1}{N}\log p\left(\boldsymbol{\theta}^N, \mathbf{x}^N\right) - \mathcal{H}\left(\Theta, \mathsf{x}\right)\right| \leqslant 2\varepsilon. \tag{2.87}$$

For any $\varepsilon > 0$ there exist an N_ε such that for all $N \geqslant N_\varepsilon$ and for any $\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}$, one has $\Pr\left\{\mathbf{x}^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right)\right\} \geqslant 1-\varepsilon$ and $\Pr\left\{\left(\boldsymbol{\Theta}^N, \mathbf{x}^N\right) \in \mathcal{A}^N_{[\Theta,\mathsf{x}]2\varepsilon}\right\} \geqslant 1-2\varepsilon$. The cardinality of the set

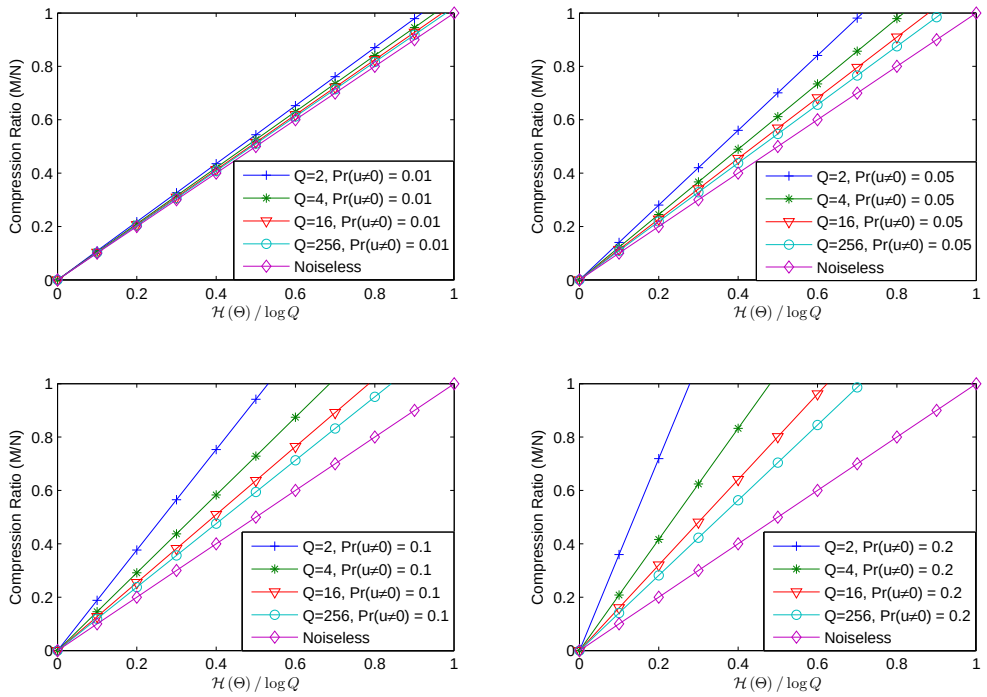

FIGURE 2.4: Optimum asymptotic achievable compression ratio in function of $\mathcal{H}(\Theta)/\log Q$, according to (2.67), for a crossover probability equal to 0.01, 0.05, 0.1, and 0.2 respectively, and without noise

$\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}$ satisfies

$$\left|\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}\right| \leqslant 2^{N(\mathcal{H}(\Theta,\mathrm{x})+2\varepsilon)}. \tag{2.88}$$

One may have ε arbitrary close to zero as $N \to \infty$.

Considering $\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}$, the estimation error probability is bounded by

$$\begin{aligned} P_\mathrm{e} &\leqslant \sum_{(\boldsymbol{\theta}^N,\mathbf{x}^N)\in\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N,\mathbf{x}^N\right\} + \sum_{(\boldsymbol{\theta}^N,\mathbf{x}^N)\notin\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) \\ &\leqslant \sum_{(\boldsymbol{\theta}^N,\mathbf{x}^N)\in\mathcal{A}^N_{[\Theta,\mathrm{x}]2\varepsilon}} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) \cdot \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N,\mathbf{x}^N\right\} + 2\varepsilon, \end{aligned} \tag{2.89}$$

Errors appear mainly due to a bad network coding matrix. By averaging over all $\mathbf{A} \in \mathbb{F}_Q^{M\times N}$, (2.89) becomes

$$P_\mathrm{e} \leqslant \sum_{\mathbf{A}\in\mathbb{F}_Q^{M\times N}} p(\mathbf{A}) \sum_{\boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]\varepsilon}} \sum_{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]\varepsilon}(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) \Pr\left\{\text{error} \mid \boldsymbol{\theta}^N,\mathbf{x}^N,\mathbf{A}\right\} + 2\varepsilon, \tag{2.90}$$

where $p(\mathbf{A}) = \Pr\{\mathbf{A} = \mathbf{A}\}$. $\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{x}^N, \mathbf{A}\right\}$ can be written as

$$\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{x}^N, \mathbf{A}\right\} = \begin{cases} 1 & \text{if } \exists\, \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\} \text{ s.t. (2.85) holds,} \\ 0 & \text{if } \forall\, \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}\text{, (2.85) does not hold.} \end{cases} \tag{2.91}$$

Using again the idea of Lemma 2.1, probability is bounded by

$$\Pr\left\{\text{error} \mid \boldsymbol{\theta}^N, \mathbf{x}^N, \mathbf{A}\right\} \leqslant \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}} \frac{\sum_{\mathbf{z}_1^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N}}{\sum_{\mathbf{z}_2^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N, \mathbf{z}_2^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_2^N}}. \tag{2.92}$$

From (2.90) and (2.92), one gets

$$P_{\mathrm{e}} \leqslant \sum_{\substack{\mathbf{A} \in \mathbb{F}_Q^{M \times N} \\ \boldsymbol{\theta}^N \in \mathcal{A}_{[\Theta]_\varepsilon}^N}} p(\mathbf{A}) \sum_{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\theta}^N, \mathbf{x}^N\right) \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}} \frac{\sum_{\mathbf{z}_1^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N}}{\sum_{\mathbf{z}_2^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N, \mathbf{z}_2^N\right) 1_{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_2^N}} + 2\varepsilon. \tag{2.93}$$

Now, for some $\boldsymbol{\theta}^N \in \mathcal{A}_{[\Theta]_\varepsilon}^N$, consider the direct image by $\mathbf{A}$ of the conditional typical set $\mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)$

$$\mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right) = \left\{\mathbf{y}^M = \mathbf{A}\mathbf{x}^N, \text{ for all } \mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)\right\}. \tag{2.94}$$

Lemma 2.3. *For any arbitrary real-valued function $h\left(\mathbf{x}^N\right)$ with $\mathbf{x}^N \in \mathbb{F}_Q^N$, one has*

$$\sum_{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)} h\left(\mathbf{x}^N\right) = \sum_{\mathbf{y}^M \in \mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)} \sum_{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)} h\left(\mathbf{x}^N\right) 1_{\mathbf{y}^M = \mathbf{A}\mathbf{x}^N}. \tag{2.95}$$

Proof. For a given $\mathbf{y}^M \in \mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)$, consider the set

$$\mathcal{X}_\varepsilon\left(\mathbf{y}^M, \mathbf{A}, \boldsymbol{\theta}^N\right) = \left\{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right) \text{ such that } \mathbf{y}^M = \mathbf{A}\mathbf{x}^N\right\}. \tag{2.96}$$

Then one has

$$\mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right) = \bigcup_{\mathbf{y}^M \in \mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)} \mathcal{X}_\varepsilon\left(\mathbf{y}^M, \mathbf{A}, \boldsymbol{\theta}^N\right), \tag{2.97}$$

with $\mathcal{X}_\varepsilon\left(\mathbf{y}_i^M, \mathbf{A}, \boldsymbol{\theta}^N\right) \cap \mathcal{X}_\varepsilon\left(\mathbf{y}_j^M, \mathbf{A}, \boldsymbol{\theta}^N\right) = \emptyset$ for any $\mathbf{y}_i^M \neq \mathbf{y}_j^M$, since the multiplication by $\mathbf{A}$ is a surjection from $\mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)$ to $\mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)$. So any sum over $\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)$ can be decomposed as

$$\begin{aligned} \sum_{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)} h\left(\mathbf{x}^N\right) &= \sum_{\mathbf{y}^M \in \mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)} \sum_{\mathbf{x}^N \in \mathcal{X}_\varepsilon\left(\mathbf{y}^M, \mathbf{A}, \boldsymbol{\theta}^N\right)} h\left(\mathbf{x}^N\right) \\ &= \sum_{\mathbf{y}^M \in \mathcal{Y}_\varepsilon\left(\mathbf{A}, \boldsymbol{\theta}^N\right)} \sum_{\mathbf{x}^N \in \mathcal{A}_{[\mathsf{x}|\Theta]_\varepsilon}^N\left(\boldsymbol{\theta}^N\right)} h\left(\mathbf{x}^N\right) 1_{\mathbf{y}^M = \mathbf{A}\mathbf{x}^N}. \end{aligned} \tag{2.98}$$

□

Applying (2.95) to (2.93), one obtains

$$\begin{aligned}
P_{\mathrm{e}} \leqslant & \sum_{\substack{\mathbf{A}\in\mathbb{F}_Q^{M\times N}\\ \boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}}} p(\mathbf{A}) \sum_{\mathbf{y}^M\in\mathcal{Y}_\varepsilon(\mathbf{A},\boldsymbol{\theta}^N)} \sum_{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{x}^N} \\
& \cdot \left(\sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}} \frac{\sum_{\mathbf{z}_1^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_1^N}}{\sum_{\mathbf{z}_2^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N,\mathbf{z}_2^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_2^N}} \right) + 2\varepsilon \\
= & \sum_{\substack{\mathbf{A}\in\mathbb{F}_Q^{M\times N}\\ \boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}}} p(\mathbf{A}) \sum_{\mathbf{y}^M\in\mathcal{Y}_\varepsilon(\mathbf{A},\boldsymbol{\theta}^N)} \left(\sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}} \sum_{\mathbf{z}_1^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) \cdot 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_1^N} \right) \\
& \cdot \left(\frac{\sum_{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{x}^N}}{\sum_{\mathbf{z}_2^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N,\mathbf{z}_2^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_2^N}} \right) + 2\varepsilon \\
\leqslant & \sum_{\substack{\mathbf{A}\in\mathbb{F}_Q^{M\times N}\\ \boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}}} p(\mathbf{A}) \sum_{\mathbf{y}^M\in\mathcal{Y}_\varepsilon(\mathbf{A},\boldsymbol{\theta}^N)} \left(\sum_{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}} \sum_{\mathbf{z}_1^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_1^N} \right) + 2\varepsilon, \qquad (2.99)
\end{aligned}$$

since we have

$$\frac{\sum_{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\theta}^N,\mathbf{x}^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{x}^N}}{\sum_{\mathbf{z}_2^N\in\mathbb{F}_Q^N} p\left(\boldsymbol{\theta}^N,\mathbf{z}_2^N\right) 1_{\mathbf{y}^M=\mathbf{A}\mathbf{z}_2^N}} \leqslant 1. \qquad (2.100)$$

The bound (2.100) is tight because for N sufficiently large, the probability of the non-typical set vanishes. Recall that $\mathbf{y}^M = \mathbf{A}\mathbf{x}^N$, even though $\mathbf{x}^N$ is not explicit in (2.99). As a vector $\mathbf{y}^M$ may correspond to several $\mathbf{x}^N$s, (2.99) is further bounded by

$$\begin{aligned}
P_{\mathrm{e}} \leqslant & \sum_{\substack{\mathbf{A}\in\mathbb{F}_Q^{M\times N}\\ \boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}}} p(\mathbf{A}) \sum_{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)} \sum_{\substack{\boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}\\ \mathbf{z}_1^N\in\mathbb{F}_Q^N}} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) 1_{\mathbf{A}\mathbf{x}^N=\mathbf{A}\mathbf{z}_1^N} + 2\varepsilon \\
\leqslant & \sum_{\substack{\boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}\\ \boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}}} \sum_{\substack{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)\\ \mathbf{z}_1^N\in\mathbb{F}_Q^N}} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) \sum_{\mathbf{A}\in\mathbb{F}_Q^{M\times N}} p(\mathbf{A}) 1_{\mathbf{A}\mathbf{x}^N=\mathbf{A}\mathbf{z}_1^N} + 2\varepsilon. \qquad (2.101)
\end{aligned}$$

Since

$$\sum_{\mathbf{A}\in\mathbb{F}_Q^{M\times N}} p(\mathbf{A}) 1_{\mathbf{A}\mathbf{x}^N=\mathbf{A}\mathbf{z}_1^N} = \Pr\left\{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N\right\}, \qquad (2.102)$$

one gets

$$P_{\mathrm{e}} \leqslant \sum_{\substack{\boldsymbol{\theta}^N\in\mathcal{A}^N_{[\Theta]_\varepsilon}\\ \boldsymbol{\varphi}^N\in\mathbb{F}_Q^N\setminus\{\boldsymbol{\theta}^N\}}} \sum_{\substack{\mathbf{x}^N\in\mathcal{A}^N_{[\mathrm{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)\\ \mathbf{z}_1^N\in\mathbb{F}_Q^N}} p\left(\boldsymbol{\varphi}^N,\mathbf{z}_1^N\right) \Pr\left\{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N\right\} + 2\varepsilon. \qquad (2.103)$$

Suppose that $\left\|\mathbf{x}^N - \mathbf{z}_1^N\right\|_0 = d$. If $d = 0$, $\Pr\left\{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N\right\}$ equals 1. Otherwise we can apply Lemma 2.2, without additive noise, $\Pr\left\{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N\right\} \leqslant f\left(d, 0; \gamma, Q, M\right)$. Depending on d being

zero or not, $P_{\mathcal{A}}$ is split as follows

$$P_{\mathrm{e}} \leqslant P_{\mathcal{A}_1} + P_{\mathcal{A}_2} + 2\varepsilon, \tag{2.104}$$

where

$$P_{\mathcal{A}_1} = \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}}} \sum_{\mathbf{z}_1^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right), \tag{2.105}$$

and

$$P_{\mathcal{A}_2} = \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}}} \sum_{\substack{\mathbf{x}^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N) \\ \mathbf{z}_1^N \in \mathbb{F}_Q^N \setminus \{\mathbf{x}^N\}}} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) \Pr\left\{\mathbf{A}\mathbf{x}^N = \mathbf{A}\mathbf{z}_1^N\right\}. \tag{2.106}$$

Lemma 2.4. *A sufficient condition for $P_{\mathcal{A}_1} \leqslant 2\varepsilon$ is that, for any pair of vectors $(\boldsymbol{\theta}^N, \boldsymbol{\varphi}^N) \in \mathcal{A}^N_{[\Theta]_\varepsilon} \times \mathcal{A}^N_{[\Theta]_\varepsilon}$ such that $\boldsymbol{\theta}^N \neq \boldsymbol{\varphi}^N$,*

$$\mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right) \cap \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\varphi}^N\right) = \emptyset. \tag{2.107}$$

Proof. Assume that (2.107) is satisfied. Changing the order of summation, (2.105) becomes

$$P_{\mathcal{A}_1} = \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N} p\left(\boldsymbol{\varphi}^N\right) \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \setminus \{\boldsymbol{\varphi}^N\} \\ \mathbf{z}_1^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)}} p\left(\mathbf{z}_1^N \mid \boldsymbol{\varphi}^N\right), \tag{2.108}$$

which can be further decomposed as $P_{\mathcal{A}_1} = P_{\mathcal{A}_{11}} + P_{\mathcal{A}_{12}}$, with

$$\begin{aligned} P_{\mathcal{A}_{11}} &= \sum_{\boldsymbol{\varphi}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}} p\left(\boldsymbol{\varphi}^N\right) \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \setminus \{\boldsymbol{\varphi}^N\} \\ \mathbf{z}_1^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)}} p\left(\mathbf{z}_1^N \mid \boldsymbol{\varphi}^N\right) \\ &\overset{(a)}{\leqslant} \sum_{\boldsymbol{\varphi}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}} p\left(\boldsymbol{\varphi}^N\right) \sum_{\mathbf{z}_1^N \in \mathbb{F}_Q^N \setminus \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\varphi}^N)} p\left(\mathbf{z}_1^N \mid \boldsymbol{\varphi}^N\right) \\ &\leqslant \sum_{\boldsymbol{\varphi}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}} p\left(\boldsymbol{\varphi}^N\right) \varepsilon \leqslant \varepsilon, \end{aligned} \tag{2.109}$$

where (a) comes from the fact that if (2.107) is satisfied, one has

$$\bigcup_{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \setminus \{\boldsymbol{\varphi}^N\}} \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right) \subseteq \mathbb{F}_Q^N \setminus \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\varphi}^N\right). \tag{2.110}$$

On the other hand,

$$\begin{aligned} P_{\mathcal{A}_{12}} &= \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \mathcal{A}^N_{[\Theta]_\varepsilon}} p\left(\boldsymbol{\varphi}^N\right) \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \mathbf{z}_1^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)}} p\left(\mathbf{z}_1^N \mid \boldsymbol{\varphi}^N\right) \\ &\leqslant \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \mathcal{A}^N_{[\Theta]_\varepsilon}} p\left(\boldsymbol{\varphi}^N\right) \leqslant \varepsilon, \end{aligned} \tag{2.111}$$

since for this part

$$\bigcup_{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}} \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right) \subseteq \mathbb{F}_Q^N. \tag{2.112}$$

From (2.109) and (2.111), Lemma 2.4 is proved. □

Now consider the term (2.106),

$$\begin{aligned} P_{\mathcal{A}_2} &= \sum_{d=1}^{N} \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}}} \sum_{\substack{\mathbf{x}^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N) \\ \mathbf{z}_1^N \in \mathbb{F}_Q^N : \left\|\mathbf{x}^N - \mathbf{z}_1^N\right\|_0 = d}} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) \cdot f\left(d, 0; \gamma, Q, M\right) \\ &\leqslant \sum_{d=1}^{\lfloor \beta N \rfloor} \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}}} \sum_{\mathbf{z}_1^N \in \mathbb{F}_Q^N} \sum_{\mathbf{x}^N \in \mathbb{F}_Q^N : \left\|\mathbf{x}^N - \mathbf{z}_1^N\right\|_0 = d} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) \cdot f\left(1, 0; \gamma, Q, M\right) \\ &\quad + \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \\ \mathbf{x}^N \in \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N)}} \sum_{\substack{\boldsymbol{\varphi}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon} \setminus \{\boldsymbol{\theta}^N\} \\ \mathbf{z}_1^N \in \mathbb{F}_Q^N}} p\left(\boldsymbol{\varphi}^N, \mathbf{z}_1^N\right) \cdot f\left(\lceil \beta N \rceil, 0; \gamma, Q, M\right), \end{aligned} \tag{2.113}$$

which is similar to (2.59) in Section 2.4.1. For N sufficient large, the condition on M/N to ensure $P_{\mathcal{A}_2}$ tends to zero as $N \to \infty$ is

$$\frac{M}{N} > \frac{\mathcal{H}\left(\Theta, \mathsf{x}\right) + \varepsilon}{\log Q - \xi}, \tag{2.114}$$

for some $\xi \in \mathbb{R}^+$. Finally, we have Proposition 2.4 to conclude the sufficient condition for reliable recovery in the NS case.

Proposition 2.4. *In the NS case, fix an arbitrary small positive real number δ, there exists $\varepsilon \in \mathbb{R}^+$, $\xi \in \mathbb{R}^+$, $N_\delta \in \mathbb{N}^+$ and $M_\varepsilon \in \mathbb{N}^+$ such that for any $N > N_\delta$ and $M > M_\varepsilon$, one has $P_e < \delta$ under MAP decoding if (2.107) and (2.114) hold. One can make both ε and ξ arbitrary close to 0 as $N \to \infty$.*

Finally, the NCS case, accounting for both additive and sensing noise, has to be considered.

Proposition 2.5. *[Sufficient condition, NCS case] Considering both additive noise and sensing noise, for N and M sufficient large and positive ε, ξ arbitrary small, the reliable recovery can be ensured under MAP decoding if*

- *the additive noise is not uniformly distributed, (2.65)*
- *there is no overlapping between any two different weakly conditional typical sets, i.e., $\mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\theta}^N\right) \cap \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}\left(\boldsymbol{\varphi}^N\right) = \emptyset$ for any two typical but different $\boldsymbol{\theta}^N$ and $\boldsymbol{\varphi}^N$,*
- *the sparsity factor satisfies the constraint in (2.66),*

the compression ratio M/N is lower bounded by

$$\frac{M}{N} > \frac{\mathcal{H}\left(\Theta, \mathsf{x}\right) + \varepsilon}{\log Q - H\left(p_{\mathsf{u}}\right) - \xi}. \tag{2.115}$$

The derivations are similar to those of Proposition 2.3 and Proposition 2.4.

2.5.2 Discussion and Numerical Results

When comparing the necessary condition in Proposition 2.1 and the sufficient condition in Proposition 2.5, an interesting fact is that $\mathcal{H}(\Theta \mid \mathsf{x}) = 0$ is a sufficient condition to have (2.107). This implies that the value of $\boldsymbol{\theta}^N$ should be fixed almost surely, as long as $\mathbf{x}^N$ is known. So, (2.107) is helpful to interpret (2.12), justifying the need for the conditional entropy $\mathcal{H}(\Theta \mid \mathsf{x})$ to tend to zero as N increases. This condition may be satisfied since $\left|\mathcal{A}^N_{[\Theta]_\varepsilon}\right| \ll \left|\mathbb{F}_Q^N\right|$ as long as $\mathcal{H}(\Theta) < \log Q$. The entropy rate $\mathcal{H}(\Theta)$ can be very small, Appendix A.2 presents a possible situation where $\mathcal{H}(\Theta) = 0$. Another implicit constraint resulting from (2.107) is

$$\sum_{\boldsymbol{\theta}^N \in \mathcal{A}^N_{[\Theta]_\varepsilon}} \mathcal{A}^N_{[\mathsf{x}|\Theta]_\varepsilon}(\boldsymbol{\theta}^N) \leqslant \left|\mathbb{F}_Q^N\right| \tag{2.116}$$

which means that

$$\mathcal{H}(\Theta, \mathsf{x}) \leqslant \log Q. \tag{2.117}$$

Consider a additive noise with $\Pr(\mathsf{u} \neq 0) = 0.1$ and the transition pmf

$$p(x_n \mid \theta_n) = \begin{cases} 1 - \Pr(\mathsf{x} \neq \Theta) & \text{if } x_n = \theta_n \\ \frac{\Pr(\mathsf{x} \neq \Theta)}{Q-1} & \text{if } x_n \in \mathbb{F}_Q^N \setminus \{\theta_n\} \end{cases}, \tag{2.118}$$

where $\Pr(\mathsf{x} \neq \Theta)$ denotes the probability of the sensing error. In Figure 2.5, the lower bound of M/N is represented as a function of $\mathcal{H}(\Theta) / \log Q$, for different values of Q and for different values of $\Pr(\mathsf{x} \neq \Theta)$.

2.6 Conclusions and future work

In this chapter we have considered a WSN where network nodes observe the components of a vector sparse in some basis. Random linear network coding is used to propagate the quantized noisy measurements in the network. MAP estimation is then used by each node of the network to estimate the sparse vector. Both asymptotically necessary and sufficient conditions of the compression ratio for reliable recovery are obtained and their convergence is also shown, even in the case of sparse network coding matrices. Several previous results have been generalized by considering a stationary and ergodic source model. Both additive noise and sensing noise have been taken into account. We have shown that the choice of the sparsity factor of the network coding matrix only depends on the additive noise. Since necessary and sufficient conditions asymptotically converge, the MAP decoder achieves the optimum lower bound of the compression ratio, which can be expressed as a function of $\mathcal{H}(\Theta, \mathsf{x})$, $H(p_\mathsf{u})$, and the alphabet size.

In this chapter, the network coding matrix was assumed to be perfectly known, without specific structure. In sensor network compressive sensing applications, the structure of the coding matrix usually depends on the structure of the network. Evaluating the impact of these constraints on the

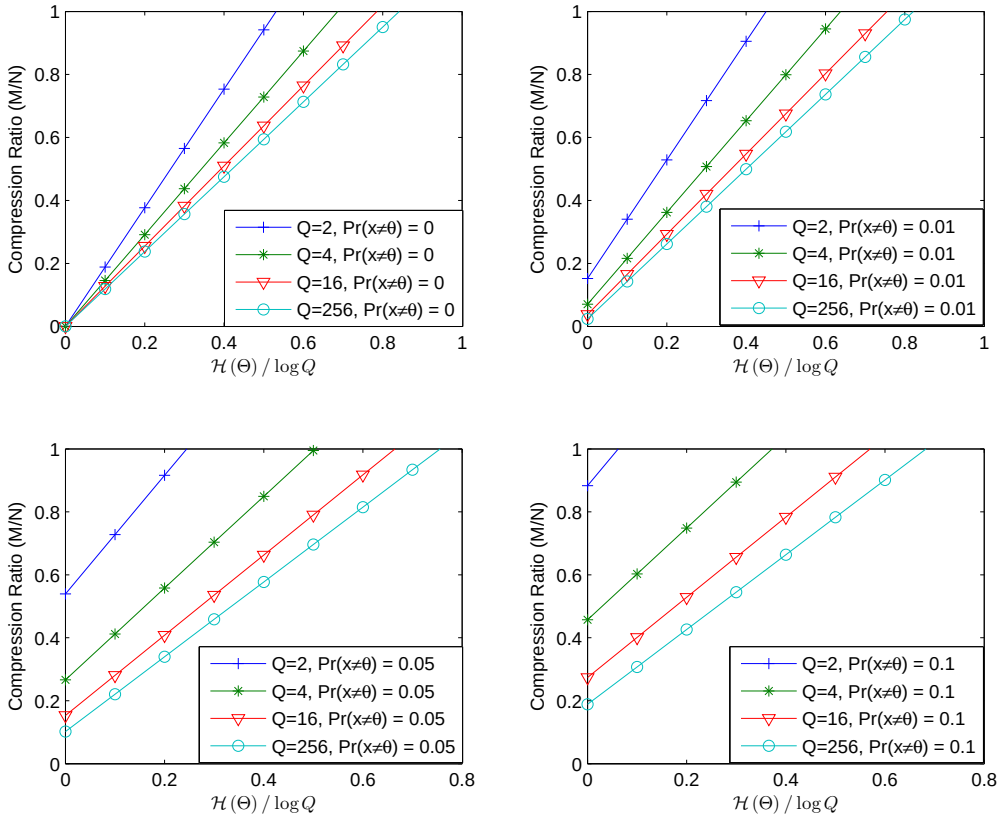

FIGURE 2.5: Optimum achievable compression ratio in function of $\mathcal{H}(\Theta) / \log Q$, according to (2.115), for the cases that $\Pr(\mathsf{x} \neq \Theta)$ being 0 (NC case), 0.01, 0.05, and 0.1, respectively, when $\Pr(\mathsf{u} \neq 0) = 0.1$

compression efficiency will be the subject of future research. A first step in this direction was done in [RAG12], which considered clustered sensors.

Chapter 3

Distributed Faulty Node Detection for WSNs

In a wireless sensor network (WSN), the sensors of some nodes may be defective and frequently produce outliers, *i.e.*, abnormal measurements. The identification of defective sensors is important to improve the global behavior of a WSN [ZMH10; MK13]. This chapter proposes and analyzes the performance of distributed algorithms to help each node of a WSN to determine whether it is equipped with a normal or a defective sensor. Several related work has been discussed in Section 1.2.2.

This chapter is organized as follows. Section 3.1 presents the system model and basic assumptions. Section 3.2 describes the two variants of the proposed DFD algorithm. Section 3.3 develops the theoretical analysis of the proposed algorithms. Section 3.4 presents some outlier models and a LODT. Section 3.5 and 3.6 provide some numerical results using simulations and experiments respectively. Section 3.7 concludes this chapter.

3.1 Notations and System Model

3.1.1 Network model

Consider an infinite plane where nodes, equipped with one sensor each, are uniformly and independently deployed, with spatial density ρ, according to a 2D homogeneous Poisson point process (PPP) [Hae12]. Let $\mathcal{S}$ denote the set of nodes. Each sensor observes some physical phenomenon and produces measurements perturbed by random noise. The noise samples at different nodes, or corrupting different measurements at the same node, are assumed statistically independent. A sensor is defective if it produces outliers. Several examples of sensing noise and outlier models are introduced in Section 3.4.1. The sets of nodes equipped with defective and non-defective (good) sensors are denoted by $\mathcal{D}$ and $\mathcal{G}$ respectively, with $\mathcal{D} \cup \mathcal{G} = \mathcal{S}$. The spatial densities of nodes with defective and good sensors are ρ_{d} and ρ_{g}, respectively, with $\rho = \rho_{\mathrm{d}} + \rho_{\mathrm{g}}$. Let θ_i denote the status (defective or good) of an arbitrary sensor node $i \in \mathcal{S}$. One has $\theta_i = 1$ if $i \in \mathcal{D}$ and $\theta_i = 0$ if $i \in \mathcal{G}$. In this work, it is assumed that θ_i remains constant in the time interval during which the DFD is performed.

Assume that any pair of nodes (i, j) can communicate only if $r_{i,j} \leqslant R_0$, where $r_{i,j}$ is their physical distance and R_0 is the communication range, dependent on the transmission power, transmitter and receiver characteristics, and propagation scenario. Define $\mathcal{N}_i = \{j \in \mathcal{S} \,|\, 0 < r_{i,j} \leqslant R_0\}$ as the set of

the neighbors of Node i. The number of nodes in $\mathcal{N}_i$ follows a Poisson distribution

$$\mathbb{P}\left\{|\mathcal{N}_i| = n\right\} = \frac{\bar{\mu}^n}{n!}\exp\left(-\bar{\mu}\right), \tag{3.1}$$

with $\bar{\mu} = \rho\pi R_0^2$, see [Hae+09]. The probability of having n_1 nodes with good sensors and $n_2 = n - n_1$ nodes with defective sensors in $\mathcal{N}_i$ is

$$\begin{aligned}\mathbb{P}\left\{|\mathcal{N}_i \cap \mathcal{G}| = n_1, |\mathcal{N}_i \cap \mathcal{D}| = n_2\right\} &= \frac{\bar{\mu}_{\mathrm{g}}^{n_1}\bar{\mu}_{\mathrm{d}}^{n_2}}{n_1!n_2!}\exp(-(\bar{\mu}_{\mathrm{g}} + \bar{\mu}_{\mathrm{d}}))\,, \\ &= P\left(\bar{\mu}_{\mathrm{g}}, \bar{\mu}_{\mathrm{d}}, n_1, n_2\right), \end{aligned} \tag{3.2}$$

where $\bar{\mu}_{\mathrm{g}} = \rho_{\mathrm{g}}\pi R_0^2$ and $\bar{\mu}_{\mathrm{d}} = \rho_{\mathrm{d}}\pi R_0^2$.

When nodes try to broadcast messages to their neighbors, we assume that some collision avoidance mechanism is put at work to limit packet losses. As a consequence, in a given finite time interval, all nodes are not necessarily able to access the channel to broadcast their message. Moreover, a message broadcast by some node is not necessarily received by all its neighbors due to transmission errors, residual packet collisions, shadowing, *etc.* Packet integrity is assumed to be determined with some CRC or checksum. Corrupted packets are considered as lost. To model all these phenomena, we assume that in the communication interval Δt, Node i only receives the messages transmitted by a subset $\mathcal{N}_i' \subseteq \mathcal{N}_i$ of its neighbors. The probability of having n nodes in $\mathcal{N}_i'$ still follows a Poisson distribution

$$\mathbb{P}\left\{\left|\mathcal{N}_i'\right| = n\right\} = \frac{\mu^n}{n!}\exp\left(-\mu\right), \tag{3.3}$$

where $\mu = \rho\pi R_0^2\left(1 - \varepsilon\left(\Delta t\right)\right)$ is the average number of nodes from which Node i receives a message during the transmission interval, see [Hae+09]. We assume that $\varepsilon\left(\Delta t\right)$ is a decreasing function of Δt accounting for all the above mentioned channel access impairments. This is reasonable, because as Δt increases more nodes have the opportunity to access the channel.

3.1.2 Local outlier detection test

Consider a generic set of nodes $\mathcal{A} \subseteq \mathcal{S}$ and define

$$\varphi\left(\mathcal{A}\right) = \begin{cases} 1, & \text{if } \mathcal{A} \cap \mathcal{D} \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \tag{3.4}$$

as the indicator function for the presence of nodes with defective sensors in $\mathcal{A}$. At a given time instant, let the random variable $M_i \in \mathbb{R}^\eta$ denote the η-dimensional data provided by the sensor of Node i, $i \in \mathcal{A}$. M_i may be a scalar or vector measurement, may contain a measurement and the value of some regressor in the case of system models linear in their parameters, or may contain measurements and experimental conditions, for general nonlinear system models.

The data provided by all nodes in $\mathcal{A}$ are gathered in $\mathbf{M}_{\mathcal{A}} = [M_i]_{i\in\mathcal{A}} \in \mathbb{R}^{\eta|\mathcal{A}|}$. For a given realization $\mathbf{m}_{\mathcal{A}} \in \mathbb{R}^{\eta|\mathcal{A}|}$ of $\mathbf{M}_{\mathcal{A}}$, denote $T(\mathbf{m}_{\mathcal{A}})$ the outcome of some LODT

$$T(\mathbf{m}_{\mathcal{A}}) = \begin{cases} 0, & \text{if no outlier is detected from } \mathbf{m}_{\mathcal{A}}, \\ 1, & \text{otherwise.} \end{cases} \tag{3.5}$$

Even if $\varphi(\mathcal{A}) = 1$, *i.e.*, at least one node with a defective sensor belongs to $\mathcal{A}$, the noise characteristics of the defective sensors may produce realizations $\mathbf{m}_{\mathcal{A}}$ that do not allow for the detection of the outliers. Thus, one introduces the probability of detection of the LODT

$$q_{\mathrm{D}}(\mathbf{M}_{\mathcal{A}}) = \mathbb{P}\left\{T(\mathbf{M}_{\mathcal{A}}) = 1 \middle| \varphi(\mathcal{A}) = 1\right\}. \tag{3.6}$$

Similarly, one considers the probability of false alarm of the LODT

$$q_{\mathrm{FA}}(\mathbf{M}_{\mathcal{A}}) = \mathbb{P}\left\{T(\mathbf{M}_{\mathcal{A}}) = 1 \middle| \varphi(\mathcal{A}) = 0\right\}. \tag{3.7}$$

In this work we consider only LODTs satisfying the following properties.

Property 3.1. *Let* $n_{\mathrm{g}} = |\mathcal{A}\cap\mathcal{G}|$ *and* $n_{\mathrm{d}} = |\mathcal{A}\cap\mathcal{D}|$, *then*

$$q_{\mathrm{FA}}(\mathbf{M}_{\mathcal{A}}) = q_{\mathrm{FA}}(n_{\mathrm{g}}), \qquad q_{\mathrm{D}}(\mathbf{M}_{\mathcal{A}}) = q_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}}). \tag{3.8}$$

Property 3.2. *Let* k *be an arbitrary node in* $\mathcal{A}$, *then*

$$\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right) = 1 \mid T(\mathbf{M}_{\mathcal{A}}) = 0, \varphi(\mathcal{A}) = 0\right\} = 0. \tag{3.9}$$

$$\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right) = 1 \mid T(\mathbf{M}_{\mathcal{A}}) = 0, \varphi(\mathcal{A}) = 1\right\} = 0. \tag{3.10}$$

In Property 3.1, we consider only the class of LODTs $T(\mathbf{M}_{\mathcal{A}})$ processing the random data vector $\mathbf{M}_{\mathcal{A}}$ in a way that disregards the knowledge of the identity of the node producing the data. For this reason, (3.6) and (3.7) depend only on the number of sensors in each status belonging to $\mathcal{A}$. In practice, the LODT takes a realization $\mathbf{m}_{\mathcal{A}}$ as input and does not need to know n_{g} and n_{d}. The notations in (3.8) indicate the dependence of the probabilities q_{FA} and q_{D} in n_{g} and n_{d} that will be exploited in the analysis of the proposed DFD algorithms involving the LODT outcomes.

Property 3.2 implies that if no outlier is detected testing the whole vector $\mathbf{m}_{\mathcal{A}}$, then no outlier will be detected testing any sub-vector of $\mathbf{m}_{\mathcal{A}}$. LODTs satisfying Properties 3.1 and 3.2 are characterized by probabilities of false alarm and of detection depending on n_{g} and n_{d} as described by the following lemma.

Lemma 3.1. *The following inequalities hold for any LODT* $T(\mathbf{M}_{\mathcal{A}})$ *satisfying Properties 3.1 and 3.2:*

$$q_{\mathrm{FA}}(n_{\mathrm{g}}) \leqslant q_{\mathrm{FA}}(n_{\mathrm{g}}+1), \tag{3.11}$$

$$q_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}}) \leqslant q_{\mathrm{D}}(n_{\mathrm{g}}+1, n_{\mathrm{d}}), \tag{3.12}$$

$$q_{\mathrm{D}}\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) \leqslant q_{\mathrm{D}}\left(n_{\mathrm{g}}, n_{\mathrm{d}}+1\right). \tag{3.13}$$

The proof of Lemma 3.1 is given in Appendix B.1. Lemma 3.1 implies that a LODT satisfying Properties 3.1 and 3.2 is more likely to detect an outlier when the number of data involved in the test increases. Examples of LODTs satisfying Properties 3.1 and 3.2 are provided in Section 3.4.2.

Properties 3.1 and 3.2 correspond to mild sufficient conditions a LODT has to satisfy to ensure the existence of an equilibrium of Algorithm 2 introduced in Section 3.2.2. LODTs, which do not satisfy these properties, may also be considered, but the behavior of Algorithm 2 cannot be analyzed theoretically in such cases.

Example 3.1. *This example introduces a LODT suited to a WSN aiming at solving a bounded-error parameter estimation problem, see [Mil+96]. In such a context, some parameter vector has to be estimated from noisy measurements. The noise corrupting the measurements provided by good sensors is assumed bounded with known bounds. Let* $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^{n_x}$ *be the vector of parameters to be estimated from the vector measurements* $\mathbf{z}_1, \ldots, \mathbf{z}_{n_S}$ *provided by* n_s *sensors. Here, the data* m_i *is a vector measurement* $\mathbf{z}_i$*. Assume that the measurement model is*

$$\mathbf{z}_i = \mathbf{z}_m\left(\mathbf{x}^*\right) + \mathbf{w}_i \tag{3.14}$$

where $\mathbf{x}^*$ *is the true value of the parameter vector,* $\mathbf{z}_m$ *is a possibly non-linear model of the measurement process, and* $\mathbf{w}_i$ *is some noise such that* $\left\|\mathbf{w}_i\right\|_\infty \leqslant \varepsilon$*, with* ε *representing some known noise bound. One may then introduce the set* $\mathbb{X}_i$ *of parameter vectors* consistent *with measurement* $\mathbf{z}_i$ *as*

$$\mathbb{X}_i = \left\{\mathbf{x} \in \mathcal{X} \middle| \left\|\mathbf{z}_i - \mathbf{z}_m\left(\mathbf{x}\right)\right\|_\infty \leqslant \varepsilon\right\} \tag{3.15}$$

and the set $\mathbb{X}$ *of parameter vectors* consistent *with* all measurements *as*

$$\mathbb{X} = \bigcap_{i=1}^{n_s} \mathbb{X}_i = \left\{\mathbf{x} \in \mathcal{X} \middle| \left\|\mathbf{z}_i - \mathbf{z}_m\left(\mathbf{x}\right)\right\|_\infty \leqslant \varepsilon, i = 1 \ldots n_s\right\}. \tag{3.16}$$

Accurate inner and outer-approximations of $\mathbb{X}$ *may be obtained, even for models* $\mathbf{z}_m$ *non-linear in* $\mathbf{x}$ *with ellipsoids, parallelotopes, zonotopes, boxes, unions of boxes, see [Mil+96; Jau+01]. When* $\mathbb{X}$ *is empty, the model* $\mathbf{z}_m$ *is either not suited to describe the system of interest, or the bounded noise property* $\left\|\mathbf{w}_i\right\|_\infty \leqslant \varepsilon$ *is not satisfied for at least one measurement,* i.e.*, there is at least one defective sensor. An empty* $\mathbb{X}$ *may be obtained even with as few as two vector measurements, providing the ability to detect outliers with very few sensor readings. A LODT may then be designed in such bounded-error parameter estimation context considering a set of measurements* $\mathbf{z}_1 \ldots \mathbf{z}_{n_S}$ *by checking whether the set* $\mathbb{X}$ *introduced in* (3.16) *is empty or not.*

3.2 DFD Algorithm

This section proposes two DFD algorithms (a single-decision variant and an iterative variant). In the single-decision variant, several measurement rounds and LODTs are performed by each node.

At the end of the last round, a *single* decision concerning the status of its sensor is taken locally by each node. In the iterative variant, a decision is taken at each round and may be updated at the next one. As will be seen in Section 3.5, the most appropriate variant depends on the ratio $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ of defective sensors in the network. In what follows, we assume that the network topology does not change within a round of both algorithms.

3.2.1 Single-decision DFD algorithm

The proposed single-decision DFD algorithm is described in Algorithm 1. The local outlier detection stage consists of two successive phases (Phase I and Phase II), alternating during L rounds. Then, the decision stage consists of a final decision phase, whose outcome is an estimate $\widehat{\theta}_i$ of the status θ_i of the sensor of Node i.

In the ℓ-th round, during Phase I, Node i collects fresh data transmitted by its neighborhood $\mathcal{N}_i$. A subset containing a proportion ζ of randomly chosen data among the received ones is then used with Node i fresh privy data to perform the LODT described in Section 3.1.2. Define $\mathcal{U}_i = \mathcal{N}_i \cup \{i\}$ and let $\mathcal{N}_i^{(\ell,\mathrm{I})} \subset \mathcal{N}_i$ be the subset of nodes from which Node i received data (excluding itself) during Phase I of Round ℓ. Consider

$$\mathcal{V}_i^{(\ell,\mathrm{I})} = S_\zeta\left(\mathcal{N}_i^{(\ell,\mathrm{I})}\right) \cup \{i\} \tag{3.17}$$

the set of node indices whose data are involved in the LODT performed by Node i at Round ℓ. The function S_ζ selects a random subset of nodes in $\mathcal{N}_i^{(\ell,\mathrm{I})}$ such that

$$\left|S_\zeta\left(\mathcal{N}_i^{(\ell,\mathrm{I})}\right)\right| = \left\lceil \zeta \left|\mathcal{N}_i^{(\ell,\mathrm{I})}\right| \right\rceil, \tag{3.18}$$

where $\left|\mathcal{N}_i^{(\ell,\mathrm{I})}\right|$ is the cardinal number of $\mathcal{N}_i^{(\ell,\mathrm{I})}$ and $\lceil\cdot\rceil$ denotes upwards rounding. The role of ζ is illustrated in Example 3.3 and further discussed in Section 3.3.4. The LODT outcome is denoted $Y_i^{(\ell)} = T(\mathbf{M}_{\mathcal{V}_i^{(\ell,\mathrm{I})}})$, which, for a given realization $\mathbf{M}_{\mathcal{V}_i^{(\ell,\mathrm{I})}} = \mathbf{m}_{\mathcal{V}_i^{(\ell,\mathrm{I})}}$, provides $y_i^{(\ell)} = T(\mathbf{m}_{\mathcal{V}_i^{(\ell,\mathrm{I})}})$. Then, during Phase II, Node i tries to broadcast $(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,\mathrm{I})})$ to indicate the other nodes which data were involved in the LODT it just performed. At the end of Phase II, Node i has received pairs $(y_j^{(\ell)}, \mathcal{V}_j^{(\ell,\mathrm{I})})$ from a subset of nodes $\mathcal{V}_i^{(\ell,\mathrm{II})} \subseteq \mathcal{U}_i$. Denote

$$\mathcal{B}_i^{(\ell)} = \left\{ j \in \mathcal{V}_i^{(\ell,\mathrm{II})} \text{ such that } i \in \mathcal{V}_j^{(\ell,\mathrm{I})} \right\}. \tag{3.19}$$

Node i then adds to a first counter z_i all $y_j^{(\ell)}$s such that $j \in \mathcal{B}_i^{(\ell)}$. Thus only LODT outcomes involving the data produced by Node i are added to z_i. Node i also accumulates in a second counter n_i the cardinal number $\left|\mathcal{B}_i^{(\ell)}\right|$ of $\mathcal{B}_i^{(\ell)}$. This counter represents the number of LODT outcomes received by Node i in which its data were involved (including its own LODT outcomes). At the end of the L rounds the value z_i/n_i is the statistics on which the decision is taken in the final decision phase.

The decision (3.22) can result in both false alarm, with probability P_{FA}, and non-detection, with probability $P_{\mathrm{ND}} = 1 - P_{\mathrm{D}}$, where P_{D} is the detection probability. The value of ζ and of the threshold γ affect the trade-off between P_{D} and P_{FA}, and have to be adjusted to meet the targeted

Algorithm 1 Single-decision DFD

1. Initialize $\widehat{\theta}_i^{(0)} = 0$, $z_i = 0$, and $n_i = 0$ for all $i \in \mathcal{S}$.
2. For each round $1 \leq \ell \leq L$:
   - Phase I, lasting $\triangle t_{\mathrm{I}}$: Node i tries to broadcast a packet containing its local data $m_i^{(\ell)}$, receives the data produced by the nodes in $\mathcal{N}_i$, randomly selects a subset $S_\zeta\left(\mathcal{N}_i^{(\ell,\mathrm{I})}\right)$ of received data and performs the test (3.5) with outcome
   $$y_i^{(\ell)} = T\left(\mathbf{m}_{\mathcal{V}_i^{(\ell,\mathrm{I})}}\right); \tag{3.20}$$
   - Phase II, lasting $\triangle t_{\mathrm{II}}$: each node broadcasts $\left(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,\mathrm{I})}\right)$ generated in Phase I and updates z_i and n_i as follows
   $$\begin{cases} z_i = z_i + \sum_{j \in \mathcal{B}_i^{(\ell)}} y_j^{(\ell)}, \\ n_i = n_i + \left|\mathcal{B}_i^{(\ell)}\right|. \end{cases} \tag{3.21}$$
3. After L rounds:
   - Decision phase: each node i estimates the status θ_i of its sensor
   $$\widehat{\theta}_i = \begin{cases} 1 \ \text{(defective)} & \text{if } z_i/n_i \geqslant \gamma, \\ 0 \ \text{(non defective)} & \text{otherwise,} \end{cases} \tag{3.22}$$
   where γ is some threshold, $0 < \gamma \leqslant 1$.

performance (see Appendix B.2 for more details). For a fixed value of ζ, when γ is close to one, P_{D} may be low. On the other hand, P_{FA} may be high for values of γ close to zero. Increasing L provides a better averaging effect in (3.22), which reduces the variance of $\widehat{\theta}_i$. Nevertheless, L cannot be taken too large, to preserve the hypothesis that the status θ_i does not vary during the whole DFD procedure.

As will be seen in Section 3.5, the single-decision DFD algorithm performs well when $\bar{\mu}_{\mathrm{d}} \ll \bar{\mu}_{\mathrm{g}}$. However, P_{FA} rapidly increases with $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. Consider the case where $\bar{\mu}_{\mathrm{d}}$ is large enough so that every non-defective sensor node has at least one defective sensor node in its neighborhood. When $\zeta = 1$, all measurements from neighbors are involved in the LODT of Node i. As a consequence, for all sensors $i \in \mathcal{S}$, the LODT outcome will be $y_i^{(\ell)} = 1$ with a high probability. Non-defective sensors are then frequently diagnosed as defective. Reducing ζ increases the chance for Node i to get an outlier-free subset of data on which the LODT can be performed and thus reduces P_{FA}. The idea of selecting only a subset of data to perform the LODT is reminiscent to GT [Che+11].

Now, when $\bar{\mu}$ is small, *i.e.*, the degree of connectivity of the network is low, it may be useful to collect measurements from the h-hop neighborhood of each node and to broadcast test outcomes to the same h-hop neighborhood. Appendix B.3 describes the updates required in Algorithm 1 to handle DFD with multi-hop data collection and LODT result dissemination.

Examples 3.2 and 3.3 illustrate the behavior of the single decision DFD algorithm in various situations.

Example 3.2. *Figure 3.1 depicts a first example WSN, where no channel access issues nor collisions are considered, thus $\varepsilon = 0$. Let k be the only node equipped with a defective sensor. Assume that both q_{FA} defined in (3.7) and q_{ND} defined in (3.6) are negligible: all the sensors in $\mathcal{U}_k$ successfully detect outliers. Moreover consider $L = 1$ (unique round) and take $\gamma = 0.7$ and $\zeta = 1$. At the end of Phase II, $\mathcal{B}_i^{(1)} = \mathcal{U}_i$, $i = 1, \ldots, k$. Moreover, $z_k^{(1)} = |\mathcal{U}_k| = 7$, $z_1^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_1| = 4$, $z_2^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_2| = 1$, and $z_3^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_3| = 0$. Since $z_k^{(1)}/|\mathcal{U}_k| = 1 > \gamma$, $z_1^{(1)}/|\mathcal{U}_1| \approx 0.67 < \gamma$, $z_2^{(1)}/|\mathcal{U}_2| = 0.25 < \gamma$, and $z_3^{(1)}/|\mathcal{U}_3| = 0 < \gamma$, only Node k determines its sensor as defective, according to (3.22), while Nodes 1, 2, and 3 diagnose their own sensor as non-defective. All decisions are thus correct.*

Example 3.3. *Figure 3.2 represents a second example of a WSN with a minimum degree of connectivity.*

Node 2 is equipped with a defective sensor. With the same assumptions as in Example 3.2, considering $L = 1$, $\varepsilon = 0$, $\gamma = 0.7$, and $\zeta = 1$, Node 1 and 2 will determine themselves as defective. Allowing two-hop data collection and the LODT outcome dissemination worsens the situation since all nodes will have the outlier produced by Node 2 in their collected data. As a consequence, all LODTs will detect the presence of an outlier, and all nodes will determine themselves as defective. This is why it is necessary to perform the LODT only on a subset of data. Consider again a two-hop data collection with $\zeta = 0.5$. In Phase I, Nodes 1 and 4 receive data from Nodes 2 and 3. Their LODTs will provide $y_1^{(1)} = 0$ with a probability 0.5 and $y_4^{(1)} = 0$ with a probability 0.5. Node 2 always uses its own data and produces $y_2^{(1)} = 1$. Node 3 receives data from Nodes 1, 2, and 4. It randomly selects $\lceil 0.5 \times 3 \rceil = 2$ data out of these 3 data and provides thus $y_1^{(1)} = 0$ with a probability 2/3. In Phase II, Node 1 will receive LODT outcomes from Nodes 2 and 3. Only outcomes such

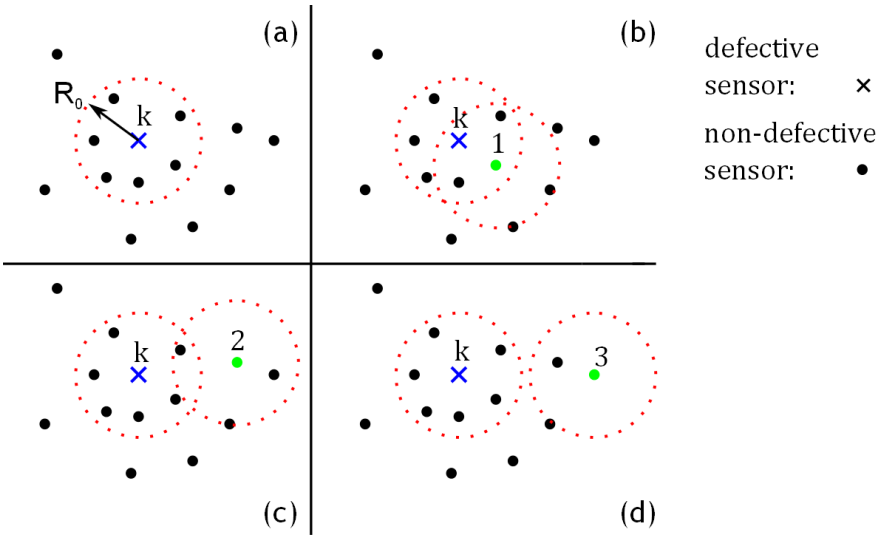

FIGURE 3.1: Example where k is the only node with a defective sensor, (a) shows that $z_k^{(1)} = |\mathcal{U}_k| = 7$. In (b), Node 1 belongs to $\mathcal{U}_k$ and $z_1^{(1)} = 4$. In (c), Node 2 is not in $\mathcal{U}_k$ but $r_{2,k} \leqslant 2R_0$, which results in $z_2^{(1)} = 1$. In (d), Node 3 has a distance $r_{3,k} > 2R_0$, so $z_3^{(1)} = 0$.

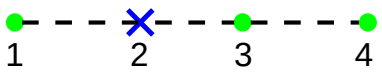

FIGURE 3.2: Example where Node 2 is the only defective sensor.

that $y_i^{(1)} = 1$ *and* $\{1\} \subset \mathcal{V}_i^{(1,I)}$ *contribute to* z_i. *Moreover, all outcomes with* $\{1\} \subset \mathcal{V}_i^{(1,I)}$ *contribute to* n_1. *The third line of Table 3.1 provides the probability vector*

$$\boldsymbol{P}_i^{(1)} = \left[\mathbb{P}\left\{\{j\} \subset \mathcal{V}_i^{(1,I)}\right\}\right]_{j=1,\ldots,4}, \ \forall i = 1,\ldots,4,$$

e.g., $\mathbb{P}\left\{\{1\} \subset \mathcal{V}_2^{(1,I)}\right\} = 2/3$. *One deduces for example*

$$\mathbb{E}\left(N_1\right) = \sum_{i=1}^{4} \mathbb{P}\left\{\{1\} \subset \mathcal{V}_i^{(1,I)}\right\} = 7/3.$$

The fourth line of Table 3.1 provides the probability vector

$$\boldsymbol{P}_i^{(2)} = \left[\mathbb{P}\left\{\{j,2\} \subset \mathcal{V}_i^{(1,I)}\right\}\right]_{j=1,\ldots,4}, \ \forall i = 1,\ldots,4,$$

which helps to evaluate $\mathbb{E}\left(Z_i\right)$, e.g.,

$$\mathbb{E}\left(Z_1\right) = \sum_{i=1}^{4} \mathbb{P}\left\{\{1,2\} \subset \mathcal{V}_i^{(1,I)}\right\} = 11/6.$$

The last line of Table 3.1 provides $\mathbb{E}\left(Z_i\right)/\mathbb{E}\left(N_i\right)$. *One observes that a proper selection of* $\gamma > 9/14$ *allows in average to detect the node producing an outlier, while avoiding false alarms. One*

Node i	1	2	3	4
$\mathbb{P}\left\{Y_i^{(1)}=1\right\}$	$1/2$	1	$2/3$	$1/2$
$\boldsymbol{P}_i^{(1)}$	$(1,\frac{1}{2},\frac{1}{2},0)$	$(\frac{2}{3},1,\frac{2}{3},\frac{2}{3})$	$(\frac{2}{3},\frac{2}{3},1,\frac{2}{3})$	$(0,\frac{1}{2},\frac{1}{2},1)$
$\boldsymbol{P}_i^{(2)}$	$(\frac{1}{2},\frac{1}{2},0,0)$	$(\frac{2}{3},1,\frac{2}{3},\frac{2}{3})$	$(\frac{1}{3},\frac{2}{3},\frac{2}{3},\frac{1}{3})$	$(0,\frac{1}{2},0,\frac{1}{2})$
$\mathbb{E}(Z_i)$	$3/2$	$8/3$	$4/3$	$3/2$
$\mathbb{E}(N_i)$	$7/3$	$8/3$	$8/3$	$7/3$
$\mathbb{E}(Z_i)/\mathbb{E}(N_i)$	$9/14$	1	$1/2$	$9/14$

TABLE 3.1: Illustration of the single decision DFD for Example 3.3

also observes that $\mathbb{E}(Z_i)/\mathbb{E}(N_i)$ is larger for Nodes 1 and 4 which are connected (at one or two hops) to fewer good nodes than Node 3. This illustrates the fact that the level of protection against a false alarm is related to the number of connections to good nodes.

Example 3.3 illustrates the importance of the choices of the values of γ and ζ. Some insights are provided for their tuning in Section 3.3.4 and Appendix B.2.

3.2.2 Iterative DFD algorithm

To improve the single-decision DFD algorithm for increasing $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$, an iterative variant of Algorithm 1 is described in Algorithm 2. The iterative algorithm is composed of two parts, of duration L_1 and L_2 rounds, respectively. In each round of the first part (Lines 1 to 4 of Algorithm 2) nodes collect data from their neighborhood, perform a LODT, and are allowed to take temporary decisions $\widehat{\theta}_i^{(\ell)}$ about the status of their sensor (*i.e.*, a single round consists of both the local outlier detection and the decision stages). The temporary decision at Round $\ell-1$ affects the set of sensor data tested by Node i at Round ℓ, *i.e.*, the nodes with $\widehat{\theta}^{(\ell-1)}=1$ remain silent during Round ℓ. The set of active neighbors of Node i is then $\widehat{\mathcal{U}}_i^{(\ell)}=\{j\in\mathcal{U}_i \text{ s.t. } \widehat{\theta}_j^{(\ell-1)}=0\}$ and one has $\mathcal{V}_i^{(\ell,\mathrm{I})}\subseteq\widehat{\mathcal{U}}_i^{(\ell)}$ and $\mathcal{V}_i^{(\ell,\mathrm{II})}\subseteq\widehat{\mathcal{U}}_i^{(\ell)}$. Any temporary decision can be updated during the next rounds of the first part. This is why z_i and n_i are reinitialized at the beginning of Phase I in each round of this first part. Finally, Algorithm 1 is performed during the second part (Lines 6 and 7 of Algorithm 2) involving only nodes with sensors deemed as non-defective.

The core procedure of the first part is described in Lines 2 and 3. One has $\widehat{\theta}_i^{(\ell)}=1$ if Node i and all neighbors which have used the data of Node i detect an outlier. Note that (3.23) is similar to (3.22) with $L=1$ and $\gamma=1$. This choice has been considered in the first part of Algorithm 2 to minimize the number of nodes with sensors wrongly deemed as defective, while having a reasonable probability of detection P_{D}. This allows then to work in the second part of Algorithm 2 with a smaller value of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. This is especially important when $\bar{\mu}_{\mathrm{d}}\geqslant 1$: a large amount of non-defective sensors may be diagnosed as defective even with $\gamma=1$. In Line 2, if $\widehat{\theta}_i^{(\ell-1)}=1$, Node i stops broadcasting its data in Round ℓ. Nevertheless, Node i still performs the LODT: this gives it a chance to become active again at Round $\ell+1$, in case it erroneously considered its sensor as defective at Round ℓ.

Algorithm 2 Iterative DFD

1. Set $\ell = 1$; $\widehat{\theta}_i^{(0)} = 0$ for all $i \in \mathcal{S}$.
2. Phase I, lasting $\triangle t_{\mathrm{I}}$: Set $z_i = n_i = 0$.
   (a) If $\widehat{\theta}_i^{(\ell-1)} = 0$, Node i tries to broadcast a packet containing its local data $m_i^{(\ell)}$, receives the data produced by the nodes in $\mathcal{N}_i^{(\ell,\mathrm{I})} \subset \mathcal{N}_i$, randomly selects a subset $S_\zeta\left(\mathcal{N}_i^{(\ell,\mathrm{I})}\right)$ of received data and performs the test (3.5) with outcome $y_i^{(\ell)} = T(\mathbf{m}_{\mathcal{V}_i^{(\ell,\mathrm{I})}})$;
   (b) Else Node i is silent, *i.e.*, it does not broadcast its data but receives data from its neighbors.
3. Phase II, lasting $\triangle t_{\mathrm{II}}$:
   (a) If $\widehat{\theta}_i^{(\ell-1)} = 0$, Node i tries broadcast $(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,\mathrm{I})})$, evaluates z_i and n_i, and performs the decision
   $$\widehat{\theta}_i^{(\ell)} = \begin{cases} 1, & \text{if } z_i/n_i = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.23}$$
   (b) Else, Node i performs the decision $\widehat{\theta}_i^{(\ell)} = y_i^{(\ell)}$.
4. $\ell = \ell + 1$.
5. If $\ell \leqslant L_1$, go to 2.
6. After round L_1:
   - each node i such that $\widehat{\theta}_i^{(L_1)} = 0$ sets $z_i = n_i = 0$.
   - each node i such that $\widehat{\theta}_i^{(L_1)} = 1$ is determined as defective. It stops broadcasting its data. It does not participate in the single-decision DFD during the following rounds.
7. Single-decision DFD is performed during L_2 rounds with threshold γ.

3.3 Analysis of the proposed DFD algorithms

In Section 3.3.1 we characterize analytically the probabilities of detection P_{D} and of false alarm P_{FA} of the proposed DFD algorithms, when the decision is taken after collecting a single data by each node in the neighborhood. These represent P_{D} and P_{FA} of the single-decision DFD algorithm for $L = 1$, or the probability of detection and false alarm on the temporary decision of the first stage of the iterative algorithm. Conditions for the existence of an equilibrium of the iterative algorithm are also evaluated in Section 3.3.4. An analysis of the traffic generated by both algorithms is then performed in Sections 3.3.3 and 3.3.6.

3.3.1 P_{D} and P_{FA} for a single round of the single-decision DFD algorithm

To lighten the notations, the round index is omitted in this section, *e.g.*, $Y_i^{(\ell)}$, $\mathcal{V}_i^{(\ell,\mathrm{I})}$ and $\mathcal{V}_i^{(\ell,\mathrm{II})}$ are replaced by Y_i, $\mathcal{V}_i^{(\mathrm{I})}$ and $\mathcal{V}_i^{(\mathrm{II})}$ respectively. Moreover, $\varepsilon\left(\triangle t_{\mathrm{I}}\right)$ and $\varepsilon\left(\triangle t_{\mathrm{II}}\right)$ are simplified to ε_{I} and $\varepsilon_{\mathrm{II}}$ respectively. Assume that Node i successfully accesses the channel during Phase I with probability α_1. One has $\alpha_1 \geqslant 1 - \varepsilon_{\mathrm{I}}$ since ε_{I} also accounts for channel impairments, such as packet collisions.

In order to perform the analysis of P_{D} and P_{FA} for a decision involving a single round of data collection from the neighborhood of Node i, one has to characterize the probability $\mathbb{P}\{Y_i = 1\}$ of the LODT (3.5). This probability depends on the numbers $N_{\mathrm{g}}^{(\mathrm{I})} = \left|\mathcal{V}_i^{(\mathrm{I})} \cap \mathcal{G}\right|$ and $N_{\mathrm{d}}^{(\mathrm{I})} = \left|\mathcal{V}_i^{(\mathrm{I})} \cap \mathcal{D}\right|$ of nodes equipped with good and defective sensors in $\mathcal{V}_i^{(\mathrm{I})}$. Define the function $h(n_{\mathrm{g}}, n_{\mathrm{d}})$ as the probability a LODT yields $Y_i = 1$, conditioned on $N_{\mathrm{g}}^{(\mathrm{I})} = n_{\mathrm{g}}$ and $N_{\mathrm{d}}^{(\mathrm{I})} = n_{\mathrm{d}}$, *i.e.*,

$$\begin{aligned} h\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) &= \mathbb{P}\left\{Y_i = 1 \mid N_{\mathrm{g}}^{(\mathrm{I})} = n_{\mathrm{g}}, N_{\mathrm{d}}^{(\mathrm{I})} = n_{\mathrm{d}}\right\} \\ &= \begin{cases} q_{\mathrm{FA}}\left(n_{\mathrm{g}}\right), & \text{if } n_{\mathrm{d}} = 0, \\ q_{\mathrm{D}}\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right), & \text{if } n_{\mathrm{d}} \neq 0. \end{cases} \end{aligned} \tag{3.24}$$

Now, $\mathbb{P}\{Y_i = 1 \mid \theta_i = 0\}$ can be expressed as a function of $\mu_{\mathrm{g}}^{(\mathrm{I})} = \alpha_2 \zeta \bar{\mu}_{\mathrm{g}}$ and $\mu_{\mathrm{d}}^{(\mathrm{I})} = \alpha_2 \zeta \bar{\mu}_{\mathrm{d}}$, respectively, where $\alpha_2 = (1 - \varepsilon_{\mathrm{I}})$ indicates the probability Node i successfully received a packet from one of its neighbors and ζ accounts for the probability a node has to select in its LODT the data produced by a neighboring node. The rounding effects have been neglected. One has thus

$$\begin{aligned} f_1\left(\mu_{\mathrm{g}}^{(\mathrm{I})}, \mu_{\mathrm{d}}^{(\mathrm{I})}\right) &= \mathbb{P}\{Y_i = 1 \mid \theta_i = 0\} \\ &\overset{(a)}{=} \sum_{n_{\mathrm{g}}=1}^{\infty} \sum_{n_{\mathrm{d}}=0}^{\infty} \mathbb{P}\{Y_i = 1 \mid N_{\mathrm{g}}^{(\mathrm{I})} = n_{\mathrm{g}}, N_{\mathrm{d}}^{(\mathrm{I})} = n_{\mathrm{d}}, \theta_i = 0\} \\ &\quad \cdot \mathbb{P}\{N_{\mathrm{g}}^{(\mathrm{I})} = n_{\mathrm{g}}, N_{\mathrm{d}}^{(\mathrm{I})} = n_{\mathrm{d}} \mid \theta_i = 0\} \\ &\overset{(b)}{=} \sum_{n_{\mathrm{g}}=1}^{\infty} \sum_{n_{\mathrm{d}}=0}^{\infty} h\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) \cdot P(\mu_{\mathrm{g}}^{(\mathrm{I})}, \mu_{\mathrm{d}}^{(\mathrm{I})}, n_{\mathrm{g}} - 1, n_{\mathrm{d}}), \end{aligned} \tag{3.25}$$

where P is the function defined in (3.2). Moreover, (a) is due to the fact that if $i \in \mathcal{V}_i^{(\mathrm{I})}$ and $\theta_i = 0$, then $N_{\mathrm{g}}^{(\mathrm{I})} \geqslant 1$, and (b) comes from Property 3.1, which states that q_{FA} and q_{D} depend only on the

number of good and defective sensors. Similarly, one may introduce

$$f_2\left(\mu_{\rm g}^{(\rm I)},\mu_{\rm d}^{(\rm I)}\right)=\mathbb{P}\left\{Y_i=1\mid\theta_i=1\right\}$$
$$=\sum_{n_{\rm g}=0}^{\infty}\sum_{n_{\rm d}=1}^{\infty}h\left(n_{\rm g},n_{\rm d}\right)\cdot P(\mu_{\rm g}^{(\rm I)},\mu_{\rm d}^{(\rm I)},n_{\rm g},n_{\rm d}-1). \tag{3.26}$$

In order to characterize $P_{\rm FA}$, introduce $N_{\rm g}^{(\rm II)}=|\mathcal{B}_i\cap\mathcal{G}|$, $N_{\rm d}^{(\rm II)}=|\mathcal{B}_i\cap\mathcal{D}|$ and the *conditional false alarm event*

$$\mathcal{E}_{i,\rm FA}^{n_{\rm g},n_{\rm d}}=\left\{\frac{\sum_{j\in\mathcal{B}_i}Y_j}{n_{\rm g}+n_{\rm d}}\geqslant\gamma\middle|N_{\rm g}^{(\rm II)}=n_{\rm g},N_{\rm d}^{(\rm II)}=n_{\rm d},\theta_i=0\right\} \tag{3.27}$$

representing, according to (3.23), the situation where Node i with non-defective sensor diagnoses it as defective, knowing that $N_{\rm g}^{(\rm II)}=n_{\rm g}\geqslant 1$ and $N_{\rm d}^{(\rm II)}=n_{\rm d}\geqslant 0$. Introducing

$$\tau_{\rm FA}\left(n_{\rm g},n_{\rm d}\right)=\mathbb{P}\left\{\mathcal{E}_{i,\rm FA}^{n_{\rm g},n_{\rm d}}\right\}, \tag{3.28}$$

then $P_{\rm FA}$ can be expressed as

$$P_{\rm FA}=\sum_{n_{\rm g}=1}^{\infty}\sum_{n_{\rm d}=0}^{\infty}\tau_{\rm FA}(n_{\rm g},n_{\rm d})\,\mathbb{P}\left\{N_{\rm g}^{(\rm II)}=n_{\rm g},N_{\rm d}^{(\rm II)}=n_{\rm d}\middle|\theta_i=0\right\}. \tag{3.29}$$

Two situations need to be considered. If Node i fails to access the channel, then $\mathcal{B}_i=\{i\}$, which means that $N_{\rm g}^{(\rm II)}=1$ and $N_{\rm d}^{(\rm II)}=0$. Otherwise $N_{\rm g}^{(\rm II)}-1$ and $N_{\rm d}^{(\rm II)}$ follow Poisson distributions. To evaluate their averages $\mu_{\rm g}^{(\rm II)}$ and $\mu_{\rm d}^{(\rm II)}$, one has to consider the probability $\alpha_3=(1-\varepsilon_{\rm I})(1-\varepsilon_{\rm II})/\alpha_1$ that Node i received a packet in Phase II from one of its neighbors, assuming that Node i actually accessed the channel in Phase I. Moreover, knowing that such packet comes from a neighbor of Node i, it contains a decision involving the data of Node i with a probability equal to ζ. Consequently, $\mu_{\rm g}^{(\rm II)}=\alpha_3\zeta\bar{\mu}_{\rm g}$ and $\mu_{\rm d}^{(\rm II)}=\alpha_3\zeta\bar{\mu}_{\rm d}$. Then

$$P_{\rm FA}=\alpha_1\widetilde{P}_{\rm FA}\left(\mu_{\rm g}^{(\rm II)},\mu_{\rm d}^{(\rm II)}\right)+(1-\alpha_1)\,f_1\left(\mu_{\rm g}^{(\rm I)},\mu_{\rm d}^{(\rm I)}\right) \tag{3.30}$$

where

$$\widetilde{P}_{\rm FA}\left(\mu_{\rm g}^{(\rm II)},\mu_{\rm d}^{(\rm II)}\right)=\sum_{n_{\rm g}=1}^{\infty}\sum_{n_{\rm d}=0}^{\infty}\tau_{\rm FA}\left(n_{\rm g},n_{\rm d}\right)P(\mu_{\rm g}^{(\rm II)},\mu_{\rm d}^{(\rm II)},n_{\rm g}-1,n_{\rm d}). \tag{3.31}$$

Similarly, consider Node i with defective sensor. The *conditional defective sensor detection event* is

$$\mathcal{E}_{i,\rm D}^{n_{\rm g},n_{\rm d}}=\left\{\frac{\sum_{j\in\mathcal{B}_i}Y_j}{n_{\rm g}+n_{\rm d}}\geqslant\gamma\middle|N_{\rm g}^{(\rm II)}=n_{\rm g},N_{\rm d}^{(\rm II)}=n_{\rm d},\theta_i=1\right\}, \tag{3.32}$$

where $n_{\mathrm{g}} \geqslant 0$ and $n_{\mathrm{d}} \geqslant 1$. Introducing

$$\tau_{\mathrm{D}}\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) = \mathbb{P}\left\{\mathcal{E}_{i,\mathrm{D}}^{n_{\mathrm{g}},n_{\mathrm{d}}}\right\}, \tag{3.33}$$

then P_{D} can be expressed as

$$P_{\mathrm{D}} = \alpha_1 \widetilde{P}_{\mathrm{D}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) + \left(1-\alpha_1\right) f_2\left(\mu_{\mathrm{g}}^{(\mathrm{I})}, \mu_{\mathrm{d}}^{(\mathrm{I})}\right), \tag{3.34}$$

where

$$\widetilde{P}_{\mathrm{D}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) = \sum_{n_{\mathrm{g}}=0}^{\infty} \sum_{n_{\mathrm{d}}=1}^{\infty} \tau_{\mathrm{D}}\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) P\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}, n_{\mathrm{g}}, n_{\mathrm{d}}-1\right). \tag{3.35}$$

Lemma 3.2. *P_{FA} and P_{D} are decreasing functions of γ.*

Proof. The monotonicity of P_{FA} and P_{D} with respect to γ comes from the fact that if $\gamma_1 > \gamma_2$, $\sum_{j\in\mathcal{B}_i} Y_j \geqslant \gamma_1\left(n_{\mathrm{g}}+n_{\mathrm{d}}\right)$ implies $\sum_{j\in\mathcal{B}_i} Y_j \geqslant \gamma_2\left(n_{\mathrm{g}}+n_{\mathrm{d}}\right)$. □

In what follows, we assume that γ is fixed, and express P_{FA} and P_{D} as functions of $\mu_{\mathrm{g}}^{(\mathrm{II})}$ and $\mu_{\mathrm{d}}^{(\mathrm{II})}$. To find a closed-form expression for P_{FA} and P_{D}, one has to evaluate the pmf of $\left(\sum_{j\in\mathcal{B}_i} Y_j \mid N_{\mathrm{g}}^{(\mathrm{II})}, N_{\mathrm{d}}^{(\mathrm{II})}, \theta_i\right)$. Consider now the event

$$\mathcal{Y}_{j,\mathrm{D}}^{n_{\mathrm{g}},n_{\mathrm{d}}} = \left\{ Y_j = 1 \middle| j \in \mathcal{B}_i, N_{\mathrm{g}}^{(\mathrm{II})} = n_{\mathrm{g}}, N_{\mathrm{d}}^{(\mathrm{II})} = n_{\mathrm{d}}, \theta_i = 1 \right\},$$

with $j \in \mathcal{B}_i$. For any $k \in \mathcal{B}_i$ with $j \neq k$, $\mathcal{Y}_{j,\mathrm{D}}^{n_{\mathrm{g}},n_{\mathrm{d}}}$ and $\mathcal{Y}_{k,\mathrm{D}}^{n_{\mathrm{g}},n_{\mathrm{d}}}$ are dependent. Their dependency comes from the fact that in general $\mathcal{N}_j \cap \mathcal{N}_k \neq \emptyset$. The pmf of $\sum_{j\in\mathcal{B}_i} Y_j$ is thus quite difficult to evaluate, since the dependency between the Y_js is not explicit. For this reason, only upper bounds of P_{FA} and P_{D} are derived in what follows.

Lemma 3.3. *Consider an arbitrary γ with $0 < \gamma \leqslant 1$, then for all $\mu_{\mathrm{g}} > 0$ and $\mu_{\mathrm{d}} \geqslant 0$, one has*

$$P_{\mathrm{FA}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) \leqslant f_1\left(4\bar{\mu}_{\mathrm{g}}, 4\bar{\mu}_{\mathrm{d}}\right), \tag{3.36}$$

$$P_{\mathrm{D}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) \leqslant f_2\left(4\bar{\mu}_{\mathrm{g}}, 4\bar{\mu}_{\mathrm{d}}\right). \tag{3.37}$$

The proof of Lemma 3.3 is given in Appendix B.5. The upper bounds in Lemma 3.3 are tight only if $\gamma \to 0$ but loose if $\gamma = 1$. Since $\gamma = 1$ in the first stage of the iterative algorithm, tighter upper bounds for $P_{\mathrm{FA}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right)$ and $P_{\mathrm{D}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right)$ are needed in this situation.

Lemma 3.4. *If $\gamma = 1$, one has*

$$P_{\mathrm{FA}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) \leqslant f_1\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right), \tag{3.38}$$

$$P_{\mathrm{D}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) \leqslant f_2\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right). \tag{3.39}$$

Proof. By definition of $\mathcal{E}_{i,\mathrm{FA}}^{n_\mathrm{g},n_\mathrm{d}}$ in (3.27), if $\gamma = 1$, one has

$$\begin{aligned}\tau_{\mathrm{FA}}\left(n_\mathrm{g}, n_\mathrm{d}\right) &= \mathbb{P}\left\{\bigcap_{j\in\mathcal{B}_i}\left\{Y_j = 1\right\}\middle|\theta_i = 0, N_\mathrm{g} = n_\mathrm{g}, N_\mathrm{d} = n_\mathrm{d}\right\}\\ &\leqslant \mathbb{P}\left\{Y_i = 1\middle|\theta_i = 0, N_\mathrm{g} = n_\mathrm{g}, N_\mathrm{d} = n_\mathrm{d}\right\}\\ &= h\left(n_\mathrm{g}, n_\mathrm{d}\right).\end{aligned} \tag{3.40}$$

Then (3.38) can be obtained from (3.25), (3.31), and (3.40). Equation (3.39) can be obtained similarly. □

3.3.2 Effects of the MAC layer on P_D and P_FA

This section focuses on the affects of the channel issues and packet losses on P_D and P_FA. Assume a perfect LODT, *i.e.*,

$$h\left(n_0, n_1\right) = \begin{cases}1, & \text{if } n_1 > 0 \text{ and } n_0 + n_1 > 1,\\ 0, & \text{otherwise,}\end{cases} \tag{3.41}$$

which means that an outliers is detected if and only if there exists at least two data and at least one outlier in the data provided by $\mathcal{V}_i^{(0)}$. With this setting, (3.25) and (3.26) have the closed forms

$$\begin{cases}f_1\left(n_\mathrm{g}, n_\mathrm{d}, \lambda^{(0)}\right) = 1 - \left(1-\lambda^{(0)}\right)^{n_\mathrm{d}},\\ f_2\left(n_\mathrm{g}, n_\mathrm{d}, \lambda^{(0)}\right) = 1 - \left(1-\lambda^{(0)}\right)^{n_\mathrm{s}-1},\end{cases} \tag{3.42}$$

where $n_\mathrm{s} = n_\mathrm{g} + n_\mathrm{d}$.

If $\theta_i = 1$ and $|\mathcal{B}_i| > 1$, one is sure that $Y_j = 1$ for any $j \in \mathcal{B}_i \setminus \{i\}$. Hence, the only situation to have $\hat{\theta}_i = 0$ is $|\mathcal{B}_i| = 1$ and $\left|\mathcal{V}_i^{(0)}\right| = 1$, knowing that Node i successfully accesses the channel during Phase 0. One needs to evaluate

$$\begin{aligned}&\Pr\left\{|\mathcal{B}_i| = 1 \text{ and } \left|\mathcal{V}_i^{(0)}\right| = 1\right\}\\ &= \sum_{n=0}^{n_\mathrm{s}-1}\binom{n_\mathrm{s}-1}{n}\bar{\lambda}^n(1-\bar{\lambda})^{n_\mathrm{s}-1-n}\left(1-\alpha_2\right)^n\left(1-\alpha_3\right)^n\\ &= \left(1-(\alpha_2+\alpha_3-\alpha_2\alpha_3)\bar{\lambda}\right)^{n_\mathrm{s}-1}.\end{aligned} \tag{3.43}$$

Define $\alpha_4 = \alpha_2 + \alpha_3 - \alpha_2\alpha_3$, then $P_\mathrm{ND} = 1 - P_\mathrm{D}$ is

$$P_\mathrm{ND} = \alpha_1\left(1-\alpha_4\bar{\lambda}\right)^{n_\mathrm{s}-1} + \left(1-\alpha_1\right)\left(1-\alpha_2\bar{\lambda}\right)^{n_\mathrm{s}-1}. \tag{3.44}$$

In the case where $\theta_i = 0$, the probability of false alarm is more complex. Nevertheless, one has

$$P_\mathrm{FA} \leq f_1\left(n_\mathrm{g}, n_\mathrm{d}, \lambda^{(0)}\right) = 1 - \left(1-\alpha_2\bar{\lambda}\right)^{n_\mathrm{d}}, \tag{3.45}$$

as $\mathcal{E}_i^{n_0,n_1} \subseteq \{Y_i = 1\}$. One may easily prove that P_{D} and the upper bound of P_{FA} are increasing functions of α_1, α_2, and α_3.

3.3.3 Traffic generated by the single-decision DFD algorithm

This analysis assumes that a single-hop data collection and LODT result dissemination is employed.

During Phase I of each round ℓ, the data $m_i^{(\ell)}$ collected by each node has to be transmitted, with its identifier, usually already present in the packet header. This type of information typically fits a single IEEE 802.15.4 payload [Ver+10].

During Phase II, $\left(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,\mathrm{I})}\right)$ has to be broadcast. One single bit is needed for $y_i^{(\ell)}$. The transmission of $\mathcal{V}_i^{(\ell,\mathrm{I})}$ is necessary if the network topology is dynamic or in presence of packet losses. This information requires $(\bar{\mu}+1)\lceil\log_2|\mathcal{S}|\rceil$ bits in average, assuming that each node can be identified with an index of $\lceil\log_2|\mathcal{S}|\rceil$ bits only. In a static network, nodes usually know the index of their neighbors. Thus, only the indexes of neighbors which do not participate to the LODT of Node i need to be transmitted. These nodes belong to $\mathcal{U}_i \setminus \mathcal{V}_i^{(\ell,\mathrm{I})}$. In this case, the required average number of bits boils down to $(1-\alpha_2\zeta)\,\bar{\mu}\lceil\log_2|\mathcal{S}|\rceil$.

Without the knowledge of $\mathcal{V}_i^{(\ell,\mathrm{I})}$, the algorithm can still be performed with somewhat degraded performance, as will be shown in Section 3.5.3.

3.3.4 Equilibrium of the iterative algorithm

Let the pair $(\theta_i, \hat{\theta}_i)$ denote the *state* of the sensor of Node i, where θ_i is its actual status and $\hat{\theta}_i$ is its estimated status. Among the four possible states, $(0,0)$ and $(1,1)$ are states resulting from a correct decision, $(0,1)$ corresponds to a false alarm, and $(1,0)$ corresponds to a non-detection. Let $\bar{\mu}_{\theta_i,\hat{\theta}_i}^{(\ell)}$ denote the density of sensors in the state $(\theta_i^{(\ell)}, \hat{\theta}_i^{(\ell)})$ at the beginning of round ℓ. The aim of this section is to characterize the evolution of $\bar{\mu}_{\theta_i,\hat{\theta}_i}^{(\ell)}$ to determine whether the iterative algorithm converges to a steady state.

Before the first round of the iterative algorithm, one has $\bar{\mu}_{00}^{(0)} = \bar{\mu}_{\mathrm{g}}$, $\bar{\mu}_{10}^{(0)} = \bar{\mu}_{\mathrm{d}}$, and $\bar{\mu}_{01}^{(0)} = \bar{\mu}_{11}^{(0)} = 0$. Note that at any round ℓ, $\bar{\mu}_{00}^{(\ell)} + \bar{\mu}_{01}^{(\ell)} = \bar{\mu}_{\mathrm{g}}$ and $\bar{\mu}_{10}^{(\ell)} + \bar{\mu}_{11}^{(\ell)} = \bar{\mu}_{\mathrm{d}}$. Since the actual status of a sensor is assumed constant during the rounds of the DFD algorithm, the only possible transitions are between states $(0,0)$ and $(0,1)$ and between states $(1,0)$ and $(1,1)$. The evolution of the densities during round ℓ is given by

$$\begin{pmatrix} \bar{\mu}_{00}^{(\ell+1)} \\ \bar{\mu}_{01}^{(\ell+1)} \\ \bar{\mu}_{10}^{(\ell+1)} \\ \bar{\mu}_{11}^{(\ell+1)} \end{pmatrix} = \begin{pmatrix} P_{0,00}^{(\ell)} & P_{0,10}^{(\ell)} & 0 & 0 \\ P_{0,01}^{(\ell)} & P_{0,11}^{(\ell)} & 0 & 0 \\ 0 & 0 & P_{1,00}^{(\ell)} & P_{1,10}^{(\ell)} \\ 0 & 0 & P_{1,01}^{(\ell)} & P_{1,11}^{(\ell)} \end{pmatrix} \begin{pmatrix} \bar{\mu}_{00}^{(\ell)} \\ \bar{\mu}_{01}^{(\ell)} \\ \bar{\mu}_{10}^{(\ell)} \\ \bar{\mu}_{11}^{(\ell)} \end{pmatrix} \tag{3.46}$$

where $P_{a,bc}^{(\ell)} = \mathbb{P}\left\{\hat{\theta}_i^{(\ell)} = c \mid \hat{\theta}_i^{(\ell-1)} = b, \theta_i = a\right\}$ is the transition probability from (a,b) to (a,c), for any $a,b,c \in \{0,1\}$. The sensors considered as defective in the previous iteration are silent, so all the transition probabilities are functions of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$. For different values of $\hat{\theta}_i^{(\ell-1)}$, $\hat{\theta}_i^{(\ell)}$ is obtained

according to Line 2 of Algorithm 2. More precisely, if $\hat{\theta}_i^{(\ell-1)} = 0$, one applies (3.23). According to (3.30) and (3.34), one has

$$\begin{cases} P_{0,01}^{(\ell)} = \alpha_1 \widetilde{P}_{\mathrm{FA}}\left(\mu_{00}^{(\ell,\mathrm{II})}, \mu_{10}^{(\ell,\mathrm{II})}\right) + (1-\alpha_1) f_1\left(\mu_{00}^{(\ell,\mathrm{I})}, \mu_{10}^{(\ell,\mathrm{I})}\right), \\ P_{0,00}^{(\ell)} = 1 - P_{0,01}^{(\ell)}, \\ P_{1,01}^{(\ell)} = \alpha_1 \widetilde{P}_{\mathrm{D}}\left(\mu_{00}^{(\ell,\mathrm{II})}, \mu_{10}^{(\ell,\mathrm{II})}\right) + (1-\alpha_1) f_2\left(\mu_{00}^{(\ell,\mathrm{I})}, \mu_{10}^{(\ell,\mathrm{I})}\right), \\ P_{1,00}^{(\ell)} = 1 - P_{1,01}^{(\ell)}, \end{cases} \tag{3.47}$$

where $\mu_{\theta_i,\hat{\theta}_i}^{(\ell,\mathrm{I})} = \alpha_2 \zeta \bar{\mu}_{\theta_i,\hat{\theta}_i}^{(\ell)}$ and $\mu_{\theta_i,\hat{\theta}_i}^{(\ell,\mathrm{II})} = \alpha_3 \zeta \bar{\mu}_{\theta_i,\hat{\theta}_i}^{(\ell)}$. If $\hat{\theta}_i^{(\ell-1)} = 1$, one has $\widehat{\theta}_i^{(\ell)} = y_i^{(\ell)}$ and

$$\begin{cases} P_{0,11}^{(\ell)} = f_1\left(\mu_{00}^{(\ell,\mathrm{I})}, \mu_{10}^{(\ell,\mathrm{I})}\right), & P_{0,10}^{(\ell)} = 1 - P_{0,11}^{(\ell)}, \\ P_{1,11}^{(\ell)} = f_2\left(\mu_{00}^{(\ell,\mathrm{I})}, \mu_{10}^{(\ell,\mathrm{I})}\right), & P_{1,10}^{(\ell)} = 1 - P_{1,11}^{(\ell)}. \end{cases} \tag{3.48}$$

For any variables $x \geq 0$ and $y \geq 0$, consider functions

$$F_1(x,y) = \alpha_1 x \left(f_1\left(\alpha_2\zeta x, \alpha_2\zeta y\right) - \widetilde{P}_{\mathrm{FA}}\left(\alpha_3\zeta x, \alpha_3\zeta y\right)\right) + \bar{\mu}_{\mathrm{g}}\left(1 - f_1(\alpha_2\zeta x, \alpha_2\zeta y)\right),$$

and

$$F_2(x,y) = \alpha_1 x \left(f_2\left(\alpha_2\zeta x, \alpha_2\zeta y\right) - \widetilde{P}_{\mathrm{D}}\left(\alpha_3\zeta x, \alpha_3\zeta y\right)\right) + \bar{\mu}_{\mathrm{d}}\left(1 - f_2(\alpha_2\zeta x, \alpha_2\zeta y)\right).$$

From (3.46-3.48), one thus obtains the reduced non-linear state equation

$$\begin{cases} \bar{\mu}_{00}^{(\ell+1)} = F_1\left(\bar{\mu}_{00}^{(\ell)}, \bar{\mu}_{10}^{(\ell)}\right), \\ \bar{\mu}_{10}^{(\ell+1)} = F_2\left(\bar{\mu}_{00}^{(\ell)}, \bar{\mu}_{10}^{(\ell)}\right), \end{cases} \tag{3.49}$$

which describes the evolution of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$ in the iterative DFD algorithm,

Let μ_{00}^* and μ_{10}^* be values at equilibrium of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$, respectively. From (3.49), one deduces that μ_{00}^* and μ_{10}^* should satisfy

$$\frac{\alpha_1 \mu_{00}^* \widetilde{P}_{\mathrm{FA}}\left(\alpha_3\zeta\mu_{00}^*, \alpha_3\zeta\mu_{10}^*\right)}{\bar{\mu}_{\mathrm{g}} - \mu_{00}^*} = 1 - \frac{\bar{\mu}_{\mathrm{g}} - \alpha_1\mu_{00}^*}{\bar{\mu}_{\mathrm{g}} - \mu_{00}^*} f_1(\alpha_2\zeta\mu_{00}^*, \alpha_2\zeta\mu_{10}^*), \tag{3.50}$$

and

$$\frac{\alpha_1 \mu_{10}^* \widetilde{P}_{\mathrm{D}}\left(\alpha_3\zeta\mu_{00}^*, \alpha_3\zeta\mu_{10}^*\right)}{\bar{\mu}_{\mathrm{d}} - \mu_{10}^*} = 1 - \frac{\bar{\mu}_{\mathrm{d}} - \alpha_1\mu_{10}^*}{\bar{\mu}_{\mathrm{d}} - \mu_{10}^*} f_2(\alpha_2\zeta\mu_{00}^*, \alpha_2\zeta\mu_{10}^*). \tag{3.51}$$

The existence property of the equilibrium point depend on the monotonicity of f_1, f_2, $\widetilde{P}_{\mathrm{FA}}$, and $\widetilde{P}_{\mathrm{D}}$.

Lemma 3.5. *$f_1(\mu_{00}, \mu_{10})$ and $f_2(\mu_{00}, \mu_{10})$ are non-decreasing functions of μ_{00} and μ_{10}.*

Proof. One has

$$\frac{\partial f_1}{\partial \mu_{00}} = \sum_{n_{\mathrm{g}}=1}^{\infty} \sum_{n_{\mathrm{d}}=0}^{\infty} \frac{\mu_{10}^{n_{\mathrm{d}}} \mu_{00}^{n_{\mathrm{g}}-1}}{n_{\mathrm{d}}!\,(n_{\mathrm{g}}-1)!} \exp\left(-\mu_{10} - \mu_{00}\right)\left(h\left(n_{\mathrm{g}}+1, n_{\mathrm{d}}\right) - h\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right)\right), \tag{3.52}$$

as shown in Appendix B.4. From Lemma 3.1, one has $h\left(n_{\mathrm{g}}, n_{\mathrm{d}}+1\right) \geqslant h\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right)$ and then $\frac{\partial f_1}{\partial \mu_{00}} \geqslant 0$. Similarly, one has $\frac{\partial f_1}{\partial \mu_{10}} \geqslant 0$, $\frac{\partial f_2}{\partial \mu_{00}} \geqslant 0$, and $\frac{\partial f_2}{\partial \mu_{10}} \geqslant 0$. □

The following lemma provides sufficient conditions to have monotone left hand-side expressions in (3.50) and (3.51).

Lemma 3.6. *Assume that the LODT (3.5) satisfies Properties 3.1 and 3.2. Assume also that for some reference Node i and some node with random index $K \in \mathcal{N}_i$ with $\theta_i = \theta_K = 0$ and $Y_i = 1$, the LODT is such that*

$$\mathbb{P}\left\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i\right\} > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\mathrm{g}}}, \tag{3.53}$$

$$q_{\mathrm{D}}\left(0, 2\right) > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\mathrm{d}}}, \tag{3.54}$$

then, for any μ_{10},

$$g_{\mathrm{FA}}\left(\mu_{00}, \mu_{10}\right) = \frac{\mu_{00} \widetilde{P}_{\mathrm{FA}}\left(\mu_{00}, \mu_{10}\right)}{\alpha_3 \zeta \bar{\mu}_{\mathrm{g}} - \mu_{00}} \tag{3.55}$$

is an increasing function of μ_{00} over $[0, \bar{\mu}_g]$ and for any μ_{00},

$$g_{\mathrm{D}}\left(\mu_{00}, \mu_{10}\right) = \frac{\mu_{10} \widetilde{P}_{\mathrm{D}}\left(\mu_{00}, \mu_{10}\right)}{\alpha_3 \zeta \bar{\mu}_{\mathrm{d}} - \mu_{10}} \tag{3.56}$$

is an increasing function of μ_{10} over $[0, \bar{\mu}_d]$.

The proof of Lemma 3.6 is in Appendix B.6.

Lemma 3.6 allows us to propose a sufficient condition on the LODT, so that existence and uniqueness of the equilibrium point can be guaranteed.

Proposition 3.1. *Assume that the considered LODT satisfies Properties 3.1 and 3.2, as well as (3.53) and (3.54), then (3.50-3.51) admits a solution, and an equilibrium of (3.49) exists.*

Proof. We study first the behavior of the right-hand side of (3.50) for a fixed value of μ_{10}. As shown in Section 3.3.1, both $\frac{\bar{\mu}_{\mathrm{g}} - \alpha_1 \mu_{00}}{\bar{\mu}_{\mathrm{g}} - \mu_{00}}$ and f_1 are positive and non-decreasing function of μ_{00}, then

$$s_1\left(\mu_{00}, \mu_{10}\right) = 1 - \frac{\bar{\mu}_{\mathrm{g}} - \alpha_1 \mu_{00}}{\bar{\mu}_{\mathrm{g}} - \mu_{00}} f_1\left(\alpha_2 \zeta \mu_{00}, \alpha_2 \zeta \mu_{10}\right) \tag{3.57}$$

is a non-increasing function of μ_{00}. If (3.53) is satisfied, the left-hand side of (3.50), $\alpha_1 g_{\mathrm{FA}}\left(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10}\right)$, is a strictly increasing function of μ_{00} over $[0, \bar{\mu}_{\mathrm{g}}]$. Moreover

$$\alpha_1 g_{\mathrm{FA}}\left(0, \alpha_3 \zeta \mu_{10}\right) = 0 < 1 - f_1\left(0, \mu_{10}\right),$$

and

$$\lim_{\mu_{00} \to \bar{\mu}_{\mathrm{g}}} \alpha_1 g_{\mathrm{FA}}\left(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10}\right) \gg \lim_{\mu_{00} \to \bar{\mu}_{\mathrm{g}}} s_1\left(\mu_{00}, \mu_{10}\right)$$

since $\lim_{\mu_{00} \to \bar{\mu}_{\mathrm{g}}} g_{\mathrm{FA}}\left(\mu_{00}, \mu_{10}\right) \to \infty$ and $s_1\left(\mu_{00}, \mu_{10}\right) \leqslant 1$. Therefore, there exists a unique intersection of the graphs of functions $\alpha_1 g_{\mathrm{FA}}\left(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10}\right)$ and $s_1\left(\mu_{00}, \mu_{10}\right)$ as μ_{00} goes from 0 to $\bar{\mu}_{\mathrm{g}}$,

for any fixed $\mu_{10} \in [0, \bar{\mu}_{\mathrm{d}}]$. The value μ_{00}^{+} of μ_{00} at which they intersect is thus a function of μ_{10}. One may write

$$\mu_{00}^{+} = \beta_1 (\mu_{10}),$$

where $\beta_1 : [0, \bar{\mu}_{\mathrm{d}}] \to [0, \bar{\mu}_{\mathrm{g}}]$. The continuity of β_1 is deduced from that of s_1 and g_{FA}.

Similarly, one may show that if (3.54) is satisfied, for any fixed μ_{00}, there exists a unique intersection of the graphs of $\alpha_1 g_{\mathrm{D}} (\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10})$ and

$$s_2 (\mu_{00}, \mu_{10}) = 1 - \frac{\bar{\mu}_{\mathrm{d}} - \alpha_1 \mu_{10}}{\bar{\mu}_{\mathrm{d}} - \mu_{10}} f_2 (\alpha_2 \zeta \mu_{00}, \alpha_2 \zeta \mu_{10}) \tag{3.58}$$

as μ_{10} varies from 0 to $\bar{\mu}_{\mathrm{d}}$. The value μ_{10}^{+} of μ_{10} at which both graphs intersect is thus a function of μ_{00}. One may write

$$\mu_{10}^{+} = \beta_2 (\mu_{00}),$$

where $\beta_2 : [0, \bar{\mu}_{\mathrm{g}}] \to [0, \bar{\mu}_{\mathrm{d}}]$. The continuity of β_2 is deduced from that of f_2 and g_{D}. Then, the equation

$$\mu_{00}^{*} = \beta_1 (\beta_2 (\mu_{00}^{*}))$$

admits at least a solution from Brouwer's fixed-point theorem, since $\beta_1 \circ \beta_2 : [0, \bar{\mu}_{\mathrm{g}}] \to [0, \bar{\mu}_{\mathrm{g}}]$ is continuous. In this way, Proposition 3.1 is proved. □

The unicity of the equilibrium requires additional monotonicity conditions on α and β, which are not easy to establish for a generic LODT.

If $\alpha_3 \zeta \bar{\mu}_{\mathrm{g}} < 4$ and $\alpha_3 \zeta \bar{\mu}_{\mathrm{d}} < 4$, one has $1 - 4/(\alpha_3 \zeta \bar{\mu}_{\mathrm{g}}) < 0$ and $1 - 4/(\alpha_3 \zeta \bar{\mu}_{\mathrm{d}}) < 0$. Then an equilibrium exists for any type of LODT. In practice, $\bar{\mu}_{\mathrm{d}}$ is usually small and (3.54) is easily satisfied. When $\alpha_3 \zeta \bar{\mu}_{\mathrm{g}} \geqslant 4$, (3.53) imposes some constraint on the LODT: considering two non-defective nodes i and k, the conditional probability $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ should be larger than $1 - 4/(\alpha_3 \zeta \bar{\mu}_{\mathrm{g}})$. The fact that Y_i and Y_k are dependent when $k \in \mathcal{B}_i$ helps to have $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ relatively large. The expression of a lower bound for $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ depends on the type of LODT (Section 3.4.2). In all cases, for large values of $\bar{\mu}_{\mathrm{g}}$, one has to choose small ζ to ensure equilibrium.

The fact that $\bar{\mu}_{\mathrm{g}}$ should not be too large may appear counter-intuitive. To explain this, assume that R_0 is large enough to allow all nodes to communicate directly with all other nodes and that $\mathcal{B}_i = \mathcal{U}_i$. At each node, in presence of a node with defective sensor, all LODTs are likely to yield $Y_i = 1$ for all nodes. All nodes will then stop communicating their data, perform a LODT at the next round with their own data only, be unable to detect an outlier, and finally turn on again. As a consequence, most of the nodes will turn on and off alternately, preventing the DFD algorithm to reach an equilibrium.

A simple way to address this issue is by reducing $\bar{\mu}_{\mathrm{g}}$ and $\bar{\mu}_{\mathrm{d}}$. This may be done with a smaller R_0, obtained by lowering the power with which nodes are transmitting data. Alternatively, one may use a small ζ, reducing the number of data involved in each LODT.

3.3.5 Local asymptotic stability of the iterative algorithm

Consider a small perturbation $\bar{\mu}_{00}^{(\ell)} = \mu_{00}^* + \delta_0^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)} = \mu_{10}^* + \delta_1^{(\ell)}$ around equilibrium (μ_{00}^*, μ_{10}^*). One may linearize (3.49) around equilibrium to get

$$\begin{pmatrix} \delta_0^{(\ell+1)} \\ \delta_1^{(\ell+1)} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \delta_0^{(\ell)} \\ \delta_1^{(\ell)} \end{pmatrix}, \tag{3.59}$$

where

$$\mathbf{A} = \begin{bmatrix} \frac{\partial F_1}{\partial \bar{\mu}_{00}}(\mu_{00}^*, \mu_{10}^*) & \frac{\partial F_1}{\partial \bar{\mu}_{10}}(\mu_{00}^*, \mu_{10}^*) \\ \frac{\partial F_2}{\partial \bar{\mu}_{00}}(\mu_{00}^*, \mu_{10}^*) & \frac{\partial F_2}{\partial \bar{\mu}_{10}}(\mu_{00}^*, \mu_{10}^*) \end{bmatrix}.$$

The linearized system (3.59) is asymptotically stable if the eigenvalues of $\mathbf{A}$ are in the unit circle. Providing stability conditions independently of the LODT is difficult. Nevertheless, for a given LODT, the local stability of (3.49) may be verified numerically. This requires an evaluation of the derivatives of f_1, f_2, P_{FA} and P_{D} with respect to $\bar{\mu}_{00}$ and $\bar{\mu}_{10}$ at equilibrium (Section 3.5.1).

3.3.6 Traffic generated by the iterative DFD algorithm

Compared to the single-decision DFD algorithm, the traffic generated by the iterative DFD algorithm is less heavy. Consider Round ℓ of the iterative algorithm, only the nodes with $\widehat{\theta}_i^{(\ell-1)} = 0$ will try to broadcast packets. As discussed in Section 3.3.4, the proportion of the active nodes is $\kappa^{(\ell)} = \frac{\bar{\mu}_{00}^{(\ell-1)} + \bar{\mu}_{10}^{(\ell-1)}}{\bar{\mu}}$. Therefore, the average number of bits per sensor to perform an L-round iterative algorithm is

- $\sum_{\ell=1}^{L} \kappa^{(\ell)} (\bar{\mu} + 1) \lceil \log_2 |\mathcal{S}| \rceil$, if the network topology is time-varying.
- $\sum_{\ell=1}^{L} \kappa^{(\ell)} (1 - \alpha_2 \zeta) \bar{\mu} \lceil \log_2 |\mathcal{S}| \rceil$, if the network is static.

3.4 Application Examples

This section presents different outlier models and an example of LODT to be used in the simulations presented in Section 3.5.

3.4.1 Outlier model

Assume that each sensor i gets a noisy observation of the same scalar physical quantity ϕ

$$x_i = \phi + w_i, \qquad \forall i \in \mathcal{S}. \tag{3.60}$$

The components w_i of the measurement noise in (3.60) are assumed to be realizations of independent random variables W_i. Here three different outlier models are considered, with different distributions of the measurement noise. When successive measurements are taken by the same sensor, we assume that the noise realizations are also independent.

- *UHV: Uniform distribution with High Variance.* The measurement noise of a good sensor $i \in \mathcal{G}$ is assumed to be uniformly distributed in the interval $[-\Delta, \Delta]$. For a defective sensor $i \in \mathcal{D}$, W_i is also uniformly distributed, but in a larger interval $[-\xi\Delta, \xi\Delta]$ with $\xi > 1$.
- *GHV: Gaussian distribution with High Variance.* The measurement noise follows a zero-mean Gaussian distribution, *i.e.*,

$$W_i \sim \begin{cases} \mathcal{N}\left(0, \sigma^2\right), & \text{if } i \in \mathcal{G}, \\ \mathcal{N}\left(0, (\xi\sigma)^2\right), & \text{if } i \in \mathcal{D}, \end{cases} \tag{3.61}$$

  where σ is the standard deviation and $\xi > 1$.
- *BGLV: Bias Gaussian distribution with Low Variance.* In this situation, we assume that $W_i \sim \mathcal{N}\left(E_i, \sigma^2\right)$ where the bias E_i is also a random variable. If $i \in \mathcal{G}$, then E_i is uniformly distributed in $[-\Delta, \Delta]$. If $i \in \mathcal{D}$, then E_i is uniformly distributed in $[-\xi\Delta, -\Delta] \cup [\Delta, \xi\Delta]$ with $\xi > 1$. Moreover, the value of E_i of each sensor is constant, *i.e.*, E_i does not vary over time.

3.4.2 Local outlier detection test

Consider some threshold $\nu \in \mathbb{R}$ and the interval $[m_i] = [m_i - \nu, m_i + \nu]$ of width 2ν centered around each measurement m_i. Consider a set of nodes $\mathcal{A} \subset \mathcal{S}$ and the intersection of all $[m_i]$s with $i \in \mathcal{A}$

$$\left[\widehat{\phi}\left(\mathbf{m}_{\mathcal{A}}\right)\right] = \bigcap_{i \in \mathcal{A}} [m_i]\,. \tag{3.62}$$

With and without presence of outliers, one is able to evaluate $\mathbb{P}\left\{\left[\widehat{\phi}\right] = \emptyset\right\}$ as a function of σ, ξ, and Δ for the three models introduced in Section 3.4.1. Equation (3.62) can be used to define a low-complexity LODT

$$T\left(\mathbf{m}_{\mathcal{A}}\right) = \begin{cases} 1, & \text{if } \left[\widehat{\phi}\left(\mathbf{m}_{\mathcal{A}}\right)\right] = \emptyset, \\ 0, & \text{else.} \end{cases} \tag{3.63}$$

The following example illustrates the behavior of the LODT in (3.63) with a small number of measurements.

Example 3.4. *Consider three sensors measuring some constant temperature, e.g., with actual value $t = 20°C$. Suppose that non-defective sensors have a bounded measurement error, e.g., $\pm 1°C$. Assume that the local measurement of the first sensor is $t_1 = 19.5°C$, and that two other sensors provide $t_2 = 20.8°C$ and $t_3 = 18.2°C$, respectively. Assuming that there is no defective sensor, and taking into account the bounded measurement noise, one deduces that $t \in \boldsymbol{t}_1 = [t_1 - 1, t_1 + 1] = [18.5, 20.5]$, $t \in \boldsymbol{t}_2 = [19.8, 21.8]$, and $t \in \boldsymbol{t}_3 = [17.2, 19.2]$. However since $\boldsymbol{t}_1 \cap \boldsymbol{t}_2 \cap \boldsymbol{t}_3 = \emptyset$, either the bounds on the measurement noise are too optimistic, or there is at least one outlier. Considering the second hypothesis, it remains difficult to determine which sensor produces an outlier, as $\boldsymbol{t}_1 \cap \boldsymbol{t}_2 \neq \emptyset$ and $\boldsymbol{t}_1 \cap \boldsymbol{t}_3 \neq \emptyset$.*

The test (3.63) allows to detect the presence of an outlier based on $\mathbf{m}_{\mathcal{A}}$ only, without having to consider the identity of nodes producing each measurement. This test does not give any insight

on the nodes with defective sensors. Thus q_{FA} and q_{D} defined in (3.7) and (3.6) only depend on n_{g} and n_{d}, where $n_{\mathrm{g}} = |\mathcal{A} \cap \mathcal{G}|$ and $n_{\mathrm{d}} = |\mathcal{A} \cap \mathcal{D}|$. Property 3.1 is thus satisfied.

$T(\mathbf{m}_{\mathcal{A}}) = 0$ implies $\left[\widehat{\phi}(\mathbf{m}_{\mathcal{A}})\right] \neq \emptyset$. As consequence, $\left[\widehat{\phi}(\mathbf{m}_{\mathcal{A}\setminus\{j\}})\right] \neq \emptyset$ for any $j \in \mathcal{A}$, since

$$\left[\widehat{\phi}(\mathbf{m}_{\mathcal{A}})\right] = \bigcap_{i\in\mathcal{A}} [m_i] \subseteq \bigcap_{i\in\mathcal{A}\setminus\{j\}} [m_i] = \left[\widehat{\phi}(\mathbf{m}_{\mathcal{A}\setminus\{j\}})\right].$$

Therefore, (3.10) holds and Property 3.2 is satisfied as well.

Hereafter we investigate the conditions introduced by Proposition 3.1.

Lemma 3.7. *Consider the LODT defined in (3.63). Consider a reference Node i and a node with random index $K \in \mathcal{N}_i$ such that $\theta_i = \theta_K = 0$, and $Y_i = 1$, then*

$$\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i\} \geqslant 1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2}. \tag{3.64}$$

See Appendix B.8 for the proof. From (3.64), one obtains that the condition (3.53) is $1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2} > 1 - \frac{4}{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}$ and thus $\alpha_3\zeta\bar{\mu}_{\mathrm{g}} < \frac{24\pi^2}{5+6\sqrt{3}\pi} \approx 6.3$.

3.5 Simulation Results

We consider a WSN of 1000 nodes randomly deployed according to a 2D PPP over a square of size 10×10 units, with $\bar{\mu} = 6$. To avoid boundary effects, only nodes in the central square area of size $(10 - 2R_0) \times (10 - 2R_0)$ units are considered in the evaluations of P_{D} and P_{FA}. The parameters of the outlier models and of the LODT are in Table 3.2.

TABLE 3.2: Parameters of the outlier models presented in Section 3.4.1 and of the LODT presented in Section 3.4.2.

	outlier model	outlier detection test
UHV	$\xi = 10$	$\nu = 2\Delta$
GHV	$\xi = 10$	$\nu = 6\sigma$
BGLV	$\xi = 10$, $\Delta = 10\sigma$	$\nu = 2\Delta$

The values of ν reported in Table 3.2 have been chosen in each case via Monte-Carlo simulation of the measurement models for $n_{\mathrm{g}} = 10$ and $n_{\mathrm{d}} = 1$. Since $q_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}})$ and $q_{\mathrm{FA}}(n_{\mathrm{g}})$ are increasing functions of n_{g}, the derived values of q_{D} and q_{FA} are valid upper bounds for all $n_{\mathrm{g}} \in [1, 10]$. Meanwhile, one uses $n_{\mathrm{d}} = 1$ so that the obtained q_{D} is the lower bound for all values of $n_{\mathrm{d}} \geqslant 1$. Figure 3.3 represents the evolutions of $q_{\mathrm{D}}(10, 1)$ and $q_{\mathrm{FA}}(10)$ as a function of $\omega = \nu/\Delta$ for the UHV and BGLV outlier models, and $\omega = \nu/(3\sigma)$ for the GHV outlier model. With the UHV outlier model, the noise corrupting data produced by a good sensor is bounded in $[-\Delta, \Delta]$, thus $q_{\mathrm{FA}} = 0$ if $\nu \geqslant 2\Delta$. With a GHV outlier model, $q_{\mathrm{FA}}(10) \leqslant 10^{-3}$ as soon as $\nu \geqslant 6\sigma$. With the BGLV model, $\nu = 2\Delta$ ensures both $q_{\mathrm{D}}(10, 1) > 0.95$ and $q_{\mathrm{FA}}(10) \leq 10^{-1}$.

In what follows, all results have been averaged over 2000 independent realizations of the WSN.

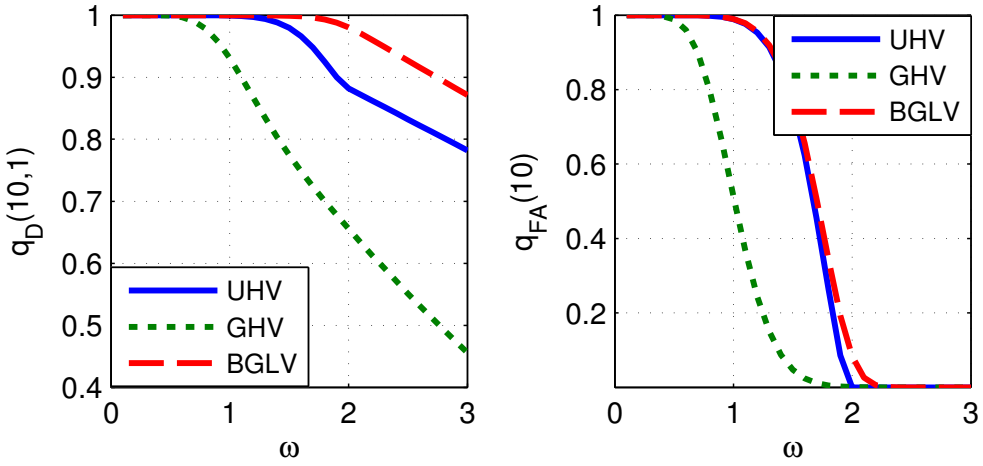

FIGURE 3.3: $q_{\mathrm{D}}(10,1)$ (left) and $q_{\mathrm{FA}}(10)$ (right) as functions of $\omega = \nu/\Delta$ for the UHV and BGLV outlier models, and $\omega = \nu/(3\sigma)$ for the GHV outlier model.

3.5.1 Performance of proposed DFD algorithm, ideal communication model

One considers here that $\varepsilon_{\mathrm{I}} = \varepsilon_{\mathrm{II}} = 0$, *i.e.*, channel access issues and packets losses are neglected.

First, the performance of a single round of the DFD algorithm described in Section 3.2.1 is evaluated. Figure 3.4 shows P_{D} as a function of P_{FA} parametrized in the threshold $\gamma \in (0,1]$, for different values of the ratio $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$ and for the different outlier models presented in Section 3.4.1. These curves are obtained with $\zeta = 1$. When γ increases, P_{D} and P_{FA} decrease in all cases, as shown by Lemma 3.2. For a given outlier model, the performance (in terms of detection accuracy) becomes worse as $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ goes larger. Among the different outlier models, the GHV provides the worst performance. This comes from the fact that nodes with defective sensors may produce measurements that appear valid with a relatively high probability. Figure 3.4 also represents with straight lines the upper bounds for P_{D} and P_{FA} provided by Lemma 3.3 for all possible values of $\gamma \in (0,1]$. For small values of γ, these bounds are relatively tight.

Figure 3.5 represents P_{D} and P_{FA} as a function of $\bar{\mu}$, with $\gamma \in \{0.8, 1\}$ and $\zeta = 1$, and for different values of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. Increasing $\bar{\mu}$ increases P_{D}, but also increases P_{FA}. For large values of $\bar{\mu}$, P_{FA} is unacceptable. Figure 3.6 represents a similar curve, but now, ζ is adapted for each value of $\bar{\mu}$ to have LODT involving less than 6 data in average. The main benefit is a decrease of P_{FA} when $\bar{\mu}$ increases. The price to be paid is a saturation of P_{D} when $\bar{\mu} > 6$.

The role played by the number of rounds on the achievable performance of the proposed DFD algorithms (single-decision and iterative) is then illustrated in Figures 3.7-3.9 corresponding to the outlier models UHV, GHV, and BGLV, respectively. P_{D} and P_{FA} are shown as functions of the round index $1 \leqslant \ell \leqslant 10$. One sets $\zeta = 1$ and $\gamma = 0.6$ in the iterative algorithm in all the cases, while the value of γ in the single-decision algorithm is adjusted so that both algorithms result to a similar P_{D} when $L = 10$. With this setting, one may evaluate the performance by comparing P_{FA} only. As can be observed, the single-decision DFD performs well when $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ is small. However, when $\bar{\mu}_{\mathrm{d}}/\bar{\mu} = 0.1$, P_{FA} is larger than 10% to keep a large P_{D}.

For the single decision algorithm, an oscillating behavior of P_{D}, depending on the iteration after

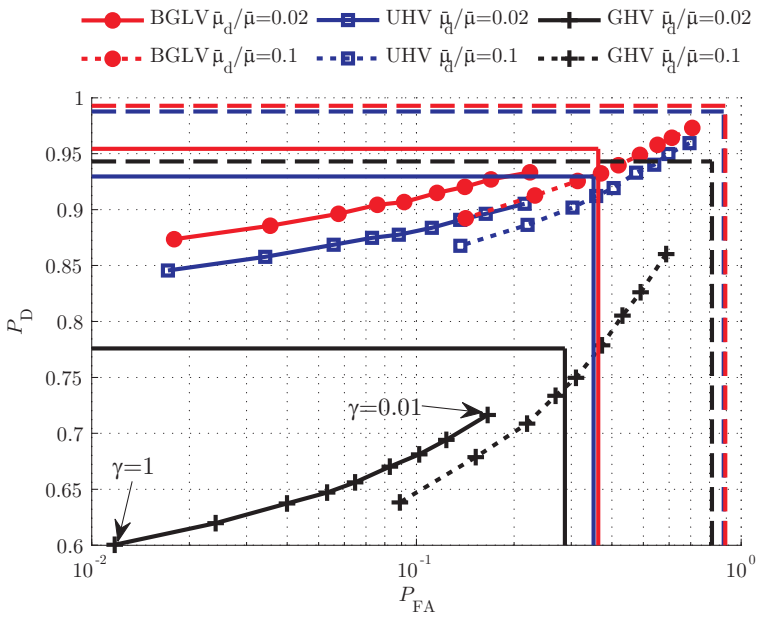

FIGURE 3.4: P_{D} as a function of P_{FA} for a single round of the DFD of Section 3.2.1 parametrized in γ, for different $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$, for $\bar{\mu} = 6$; upper bounds for P_{D} and P_{FA} are also provided (straight lines)

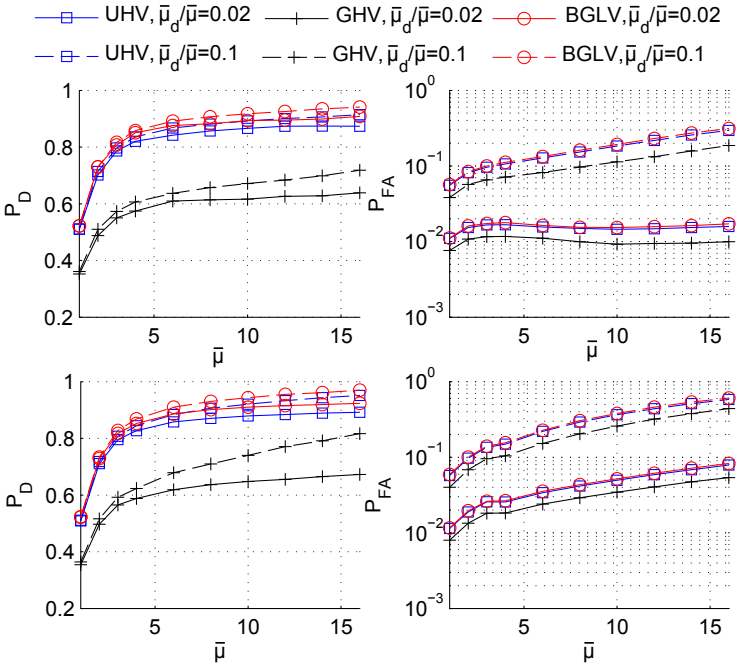

FIGURE 3.5: P_{D} (left) and P_{FA} (right) as functions of $\bar{\mu}$ for a single round of DFD for different values of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ with $\gamma \in \{0.8, 1\}$ and $\zeta = 1$.

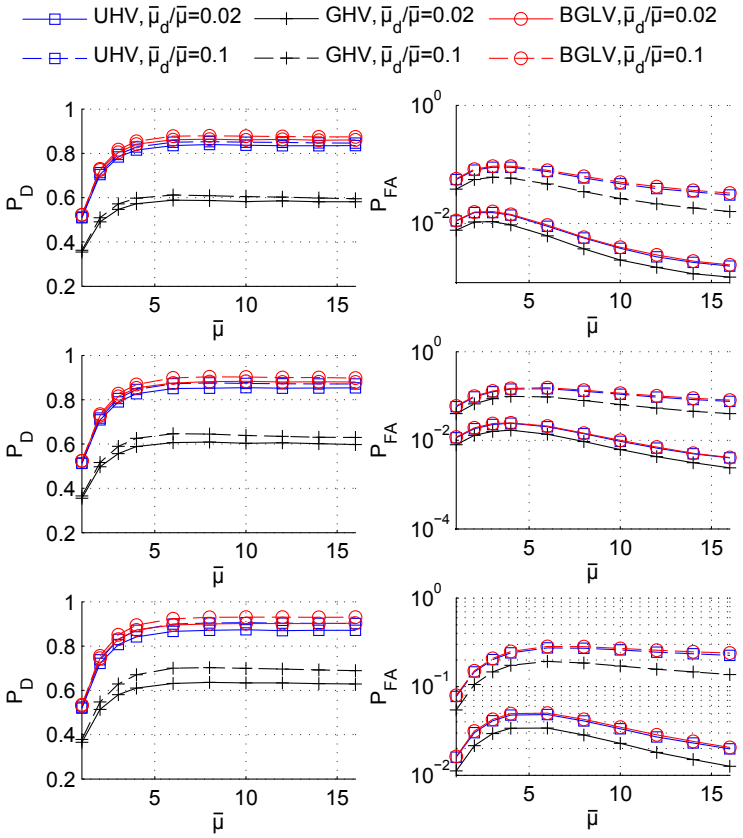

FIGURE 3.6: P_{D} (left) and P_{FA} (right) as functions of $\bar{\mu}$ for a single round of DFD for different values of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ with $\gamma \in \{0.6, 0.8, 1\}$ and ζ adapted to have LODT involving less than 6 neighboring data.

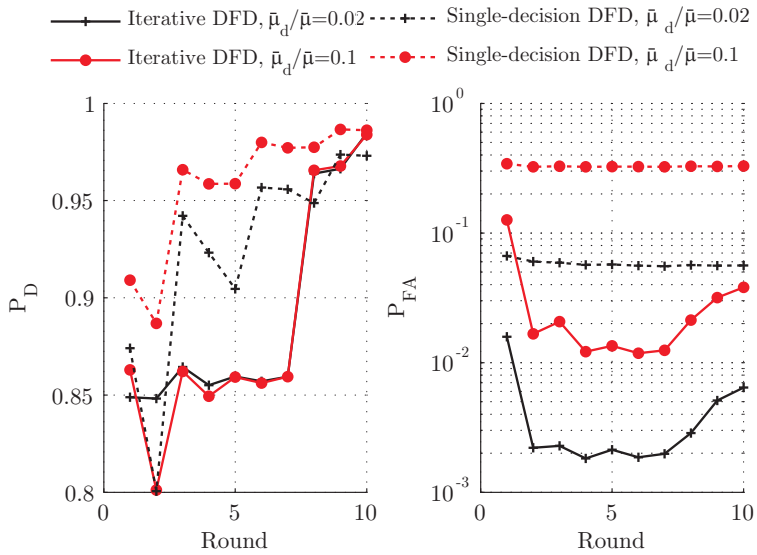

FIGURE 3.7: P_{D} (left) and P_{FA} (right) as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV; For the single-decision algorithm, $\gamma = 0.65$, for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

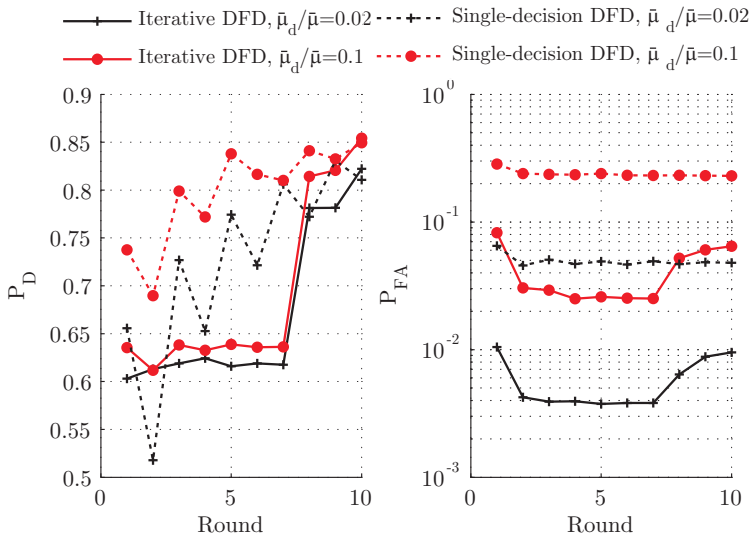

FIGURE 3.8: P_{D} (left) and P_{FA} (right) as a function of the number of rounds, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is GHV. For the single-decision algorithm, $\gamma = 0.5$ as $\bar{\mu}_{\mathrm{d}}/\bar{\mu} = 0.02$ and $\gamma = 0.59$ as $\bar{\mu}_{\mathrm{d}}/\bar{\mu} = 0.1$. For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

which the decision is taken can be observed. This is due to the decision rule (3.22), where z_i/n_i, for small values of L, may vary significantly, since n_i and z_i are small. This effect vanishes as L increases, since n_i and z_i increase as well.

For the iterative algorithm, one considers $L_1 = 7$, and $L_2 = 3$. As can be observed, the seven iterative rounds of the first stage significantly reduces P_{FA}. The oscillations of P_{D} and P_{FA} are due to the fact that, during the iterative phase of the algorithm, a decreasing proportion of nodes in false alarm alternatively turns off and on in subsequent rounds until the algorithm switches to the single-decision phase. During the following three rounds, P_{D} improves rapidly, while P_{FA} remains small. This is due to the fact that most of the defective sensors have been detected and turned off during the previous iterative phase. Hence at the beginning of the single-decision rounds the residual $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ is much lower than the initial $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. Thus the iterative DFD algorithm performs better than the single-decision variant. For the same value of P_{D}, P_{FA} is almost one order of magnitude smaller with the iterative variant.

To better understand the oscillating behavior observed in Figures 3.7-3.9, one has to study the stability of the iterative DFD algorithm. Consider for example $\bar{\mu} = 6$ and the UHV outlier model with $\xi = 10$ and $\nu = 2\Delta$. Evaluating f_1, f_2, P_{FA} and P_{D} as a function of $\bar{\mu}_{00}$ and $\bar{\mu}_{10}$, the equilibrium can be characterized numerically. Consider $\zeta = 1$ and

$$\begin{cases} E_1(\mu_{00}, \mu_{10}) = \alpha_1 g_{\mathrm{FA}}(\alpha_3\mu_{00}, \alpha_3\mu_{10}) - s_1(\alpha_3\mu_{00}, \alpha_3\mu_{10}), \\ E_2(\mu_{00}, \mu_{10}) = \alpha_1 g_{\mathrm{D}}(\alpha_3\mu_{00}, \alpha_3\mu_{10}) - s_2(\alpha_3\mu_{00}, \alpha_3\mu_{10}), \end{cases}$$

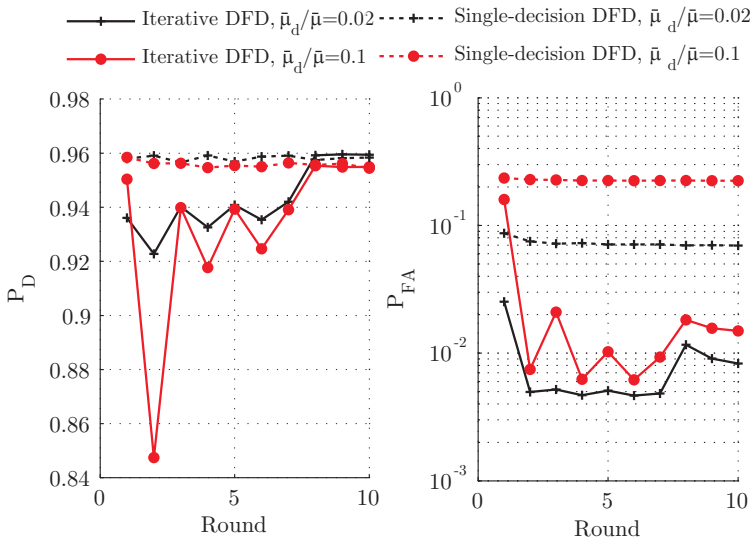

FIGURE 3.9: P_{D} (left) and P_{FA} (right) as a function of the number of rounds, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is BGLV. For the single-decision algorithm, $\gamma = 0.68$ as $\bar{\mu}_{\mathrm{d}}/\bar{\mu} = 0.02$ and $\gamma = 0.85$ as $\bar{\mu}_{\mathrm{d}}/\bar{\mu} = 0.1$ For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

TABLE 3.3: Entries of the matrix $\mathbf{A}$ in (3.59) and of its eigenvalues λ_1 and λ_2 for the UHV outlier model when $\bar{\mu} = 6$, $\xi = 10$, and $\nu = 2\Delta$

$\bar{\mu}_{\mathrm{d}}/\bar{\mu}$	0.02	0.1
$\left(\frac{\mu^*_{00}}{\bar{\mu}_{\mathrm{g}}}, \frac{\mu^*_{10}}{\bar{\mu}_{\mathrm{d}}}\right)_{\mathrm{t}}$	(0.998, 0.145)	(0.988, 0.145)
$\left(\frac{\mu^*_{00}}{\bar{\mu}_{\mathrm{g}}}, \frac{\mu^*_{10}}{\bar{\mu}_{\mathrm{d}}}\right)_{\mathrm{s}}$	(0.995, 0.142)	(0.988, 0.142)
$\mathbf{A}$	$\begin{pmatrix} 0.011, & -2.02 \\ -0.001, & -0.006 \end{pmatrix}$	$\begin{pmatrix} 0.075 & -2.01 \\ -0.005 & -0.052 \end{pmatrix}$
(λ_1, λ_2)	(0.05, 0.046)	(0.13, 0.10)

deduced from (3.50-3.51), the equilibrium conditions may be rewritten as

$$\begin{cases} E_1\left(\mu^*_{00}, \mu^*_{10}\right) = 0, \\ E_2\left(\mu^*_{00}, \mu^*_{10}\right) = 0. \end{cases}$$

Figure 3.10 shows that $E_1\left(\mu^*_{00}, \mu^*_{10}\right) = E_2\left(\mu^*_{00}, \mu^*_{10}\right) = 0$ for the pairs (μ^*_{00}, μ^*_{10}) reported in Table 3.3, with $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1, 0.2\}$. The simulation results $(\mu^*_{00}/\bar{\mu}_{\mathrm{g}}, \mu^*_{10}/\bar{\mu}_{\mathrm{d}})_{\mathrm{s}}$ obtained from Figure 3.7 are also presented in Table 3.3 showing a very good match with the theoretical results $(\mu^*_{00}/\bar{\mu}_{\mathrm{g}}, \mu^*_{10}/\bar{\mu}_{\mathrm{d}})_{\mathrm{t}}$ obtained by solving (3.50-3.51).

The derivatives of f_1, f_2, P_{FA} and P_{D} with respect to μ_{00} and μ_{10} allow to get the matrix $\mathbf{A}$ of the linearized model (3.59) and to evaluate its eigenvalues λ_1 and λ_2. Table 3.3 shows that for three different values of the ratio $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$, the eigenvalues are within the unit circle. The linearized system is thus asymptotically stable. Moreover, the norms of the eigenvalues increase with $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. The iterative DFD algorithm converges thus faster to equilibrium when $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$ is small.

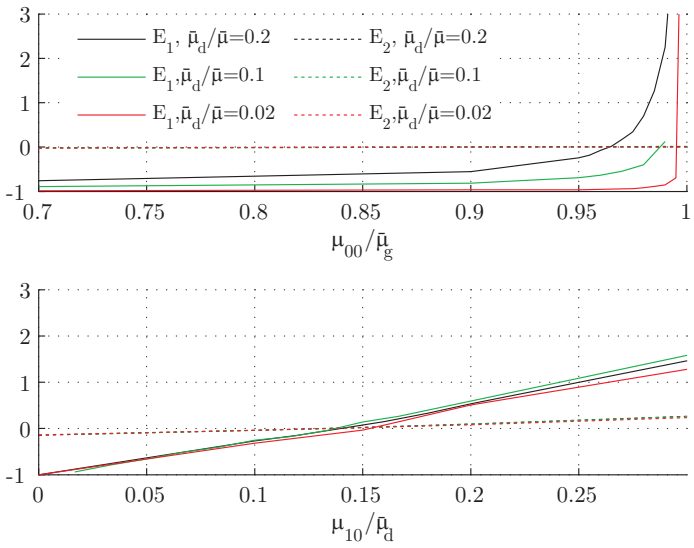

FIGURE 3.10: Evolution of $E_1(\mu_{00}, \mu_{10}^*)$ and $E_2(\mu_{00}, \mu_{10}^*)$ as functions of $\mu_{00}/\bar{\mu}_g$ (top), and of $E_1(\mu_{00}^*, \mu_{10})$ and $E_2(\mu_{00}^*, \mu_{10})$ as functions of $\mu_{10}/\bar{\mu}_d$ (bottom), with $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1, 0.2\}$ and $\zeta = 1$

3.5.2 Performance of proposed DFD algorithm, realistic communication model

In this section, channel access issues as well as packet losses are considered. For that purpose, we assume $\alpha_1 = 0.95$, $\varepsilon_I = \varepsilon_{II} = 0.10$. This leads to $\alpha_2 \approx 0.9$, and $\alpha_3 \approx 0.85$. A node manages to access the channel in 95% of the cases during a given phase of the considered algorithms. It has 90% chance to receive in Phase I a packet from one of its neighbors, and 85% chance to receive in Phase II a packet containing a decision involving the data it manages to broadcast in Phase I.

Figure 3.11 is the counterpart of Figure 3.7 for the UHV outlier model. P_D (left) and P_{FA} (right) are depicted as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. For the single-decision algorithm, $\gamma = 0.65$, whereas for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

As far as the iterative algorithm is concerned, the values of P_D and P_{FA} are very close to those obtained with ideal communication conditions. During iterations, the oscillations are better damped. For example, when $\bar{\mu}_d/\bar{\mu} = 0.1$, it results $\lambda_1 = 0.034$ and $\lambda_2 = -0.117$.

For the single decision algorithm, performance are more degraded, *e.g.*, P_D is reduced from 0.97 to 0.91 when the more realistic communication model is considered.

3.5.3 Comparison with other DFD solutions

This section compares the proposed approach with two alternative DFD algorithms presented in [CKS06] and in [LC08]. These reference DFD schemes have a LODT relatively close to that considered here. Nevertheless, a fair comparison between the two algorithms is quite difficult because the tests have some differences in their form and parameters.

First, ideal communication conditions are considered. In [CKS06], in the first stage, each node

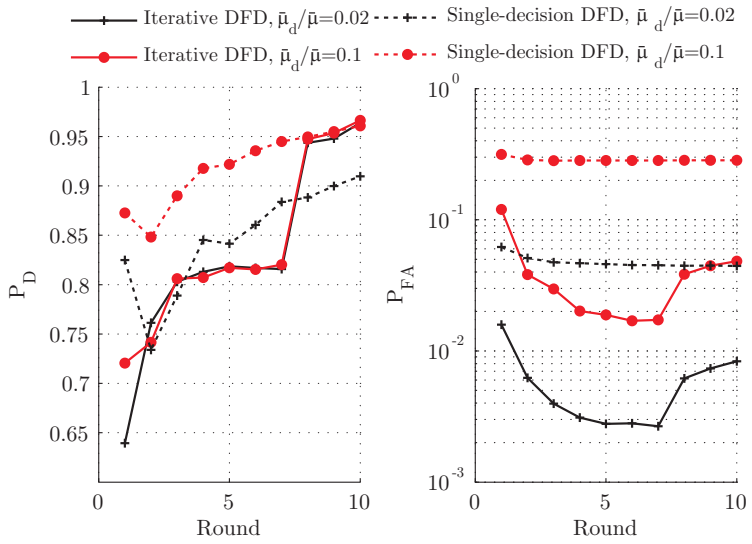

FIGURE 3.11: P_{D} (left) and P_{FA} (right) as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_{\mathrm{d}}/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV; For the single-decision algorithm, $\gamma = 0.65$, for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

exchanges its measurement with its neighbors and then performs a LODT. Compared to the proposed approach, Node i performs pairwise comparisons of its local measurement with all measurements received from its neighbors, to produce a binary value $c_{i,j}$ indicating whether there exists an outlier among m_i and m_j. Then, Node i decides whether its sensor is defective based on $c_{i,j}$, $j \in \mathcal{N}_i$. In a second stage, Node i broadcasts the estimate of its status, as well as its table of neighbors (including the index of Node j and $c_{i,j}$), which needs $(\bar{\mu}+1)\,(1+\log_2|\mathcal{S}|)$ bits in average. This is of the same order of magnitude as the results indicated in Sections 3.3.3 and 3.3.6.

Figure 3.12 compares the proposed DFD and the DFD algorithm in [CKS06] in terms of P_{D} and P_{FA}, after 10 rounds of the algorithms, with $\bar{\mu} = 6$. In fact, the only parameter one can control in the DFD algorithm of [CKS06] is the threshold of a local test (similar to ν in our case), whereas the performance of the proposed DFD depends on ν, γ, ζ, L_1, and L_2. For example, we fix ν (as the value in Table 3.2), $L_1 = 7$, $L_2 = 3$ and choose different values of γ and $\zeta = 1$. As can be seen in Figures 3.12, the $(P_{\mathrm{D}}, P_{\mathrm{FA}})$ curve of the proposed DFD algorithm is in most of the cases above that of the DFD algorithm of [CKS06].

Now packet losses and the channel access issues are taken into account. The DFD algorithm of [LC08] is considered, since it is a modified version of that of [CKS06] accounting for channel impairments. The UHV and BGLV models are considered for different values of the probability P_{success} of successful channel access and transmission. Results shown in Figure 3.13 compare also the performance of the proposed algorithm with and without transmission of nodes indices in $\mathcal{V}_i^{(\ell,\mathrm{I})}$. The performance of the algorithm in [LC08] is also indicated. With or without the knowledge of $\mathcal{V}_i^{(\ell,\mathrm{I})}$, the performance of the proposed algorithm are very close, especially when $P_{\mathrm{success}} \geqslant 0.8$. The amount of data that needs to be transmitted between nodes can thus be reduced to one bit per node for the LODT result dissemination phase.

For the UHV model, the DFD algorithm in [LC08] provides lower P_{FA} for a similar P_{D} than

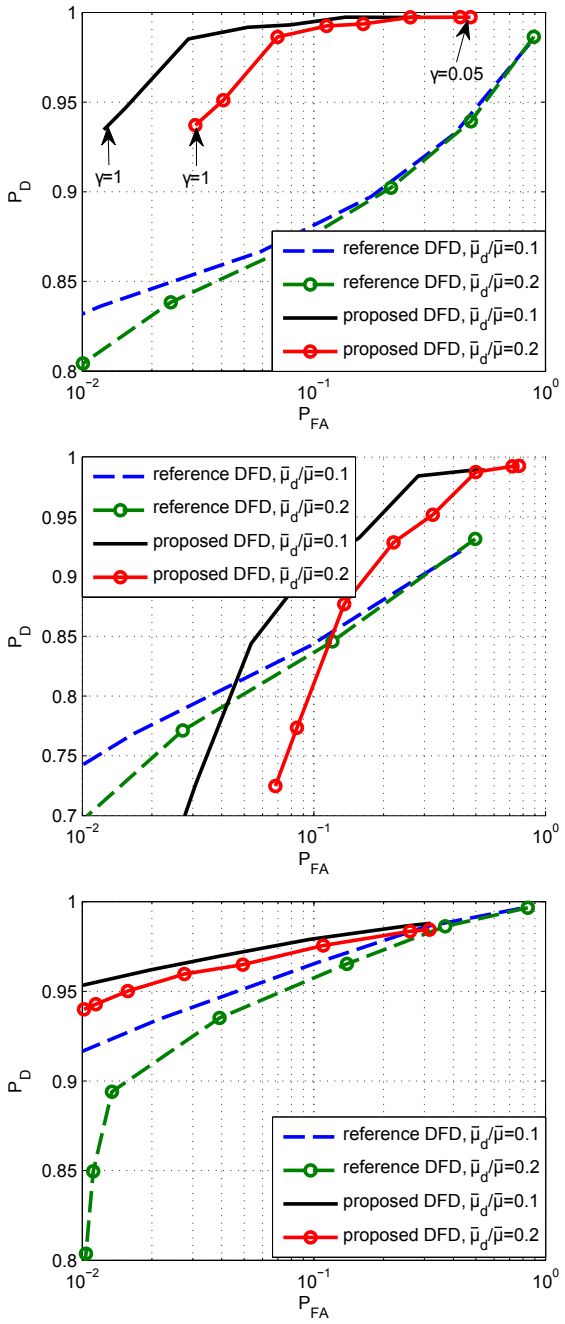

FIGURE 3.12: P_{D} as a function of P_{FA} for the DFD algorithm in [CKS06] and the proposed DFD after 10 rounds, with $\bar{\mu}_d/\bar{\mu} \in \{0.1, 0.2\}$ and the outlier models UHV (top), GHV (middle), and BGLV (bottom).

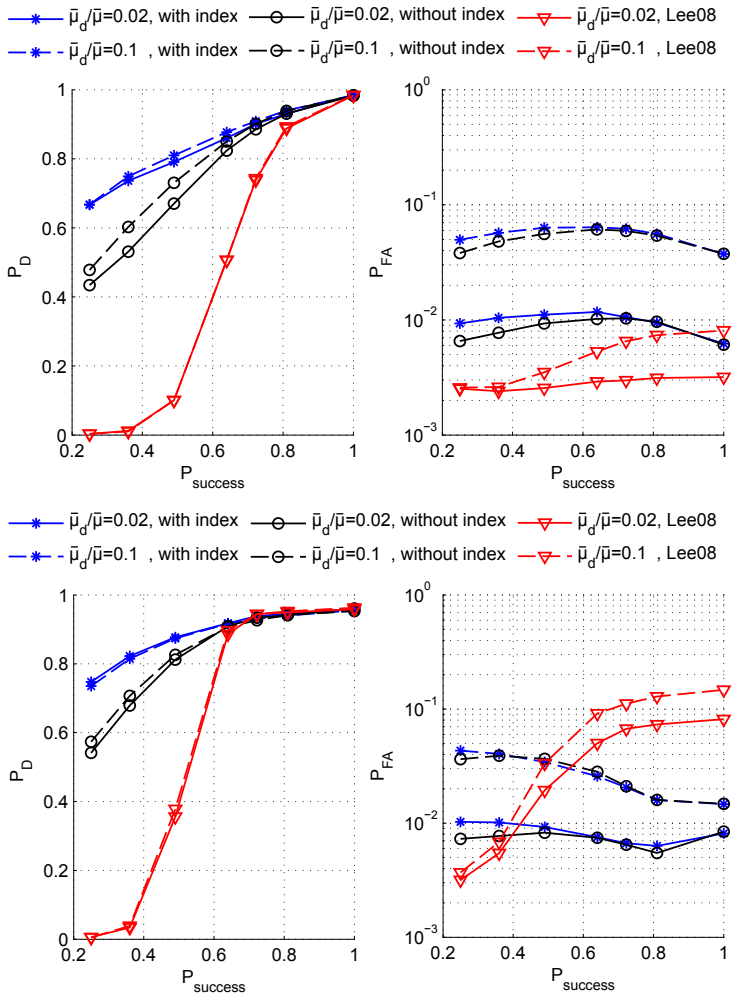

FIGURE 3.13: Comparison of the iterative algorithms with or without the transmission of $\mathcal{V}_j^{(\ell,\mathrm{I})}$, as well as the DFD algorithm of [LC08], with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_\mathrm{d}/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV (left) and BGLV (right). For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

the proposed algorithm only when P_{success} is close to 1. When P_{success} decreases the price to be paid in [LC08] for a low P_{FA} translates in a significantly degraded P_{D}. For the BGLV model, the DFD algorithm in [LC08] provides higher P_{FA} for a similar P_{D} for $P_{\text{success}} \in [0.6, 1]$ compared to the proposed approach. When $P_{\text{success}} < 0.6$, the value of P_{D} significantly decreases, even if P_{FA} decreases too. In this case, the proposed approach performs better for all values of P_{success}.

3.6 Experimental Results

To further investigate the impact of the protocol stack and of real propagation conditions, the proposed algorithm has also been implemented on the Data Sensing and Processing Testbed (DataSens), which is a part of the experimental facilities available within EuWIn@CNIT/Bologna. The testbed we adopted consisted of 20 or 41 wireless sensor nodes of type EMB-2530PA [Ins]. The TIMAC software stack [Tim], compliant with the IEEE 802.15.4 standard, has been used for the realization

of the algorithm. The obtained results are then compared with the simulation results to evaluate effectiveness of the algorithm. The proposed algorithm achieves good performance (both NDR and FAR are under 5%) within a short time (1s), even under an unstable communication environment.

FIGURE 3.14: EMB-2530PA with debugger.

3.6.1 Modified DFD for implementation

Figure 3.15 presents a modified version of Algorithm 2. Recall that L_1 and L_2 represent the number of iterative rounds and of non-iterative rounds, respectively. Each round ℓ involves two transmission phases of identical duration Δt indexed by $k \in \{0, 1\}$. Notice that $p_i^{(\ell,k)}$ represents the MAC payload that needs to be transmitted during Phase k of Round ℓ. According to Algorithm 2, one has $p_i^{(\ell,0)} = m_i^{(\ell)}$ and $p_i^{(\ell,1)} = (y_i^{(\ell)}, \mathcal{V}_i^{(\ell,0)})$.

To improve the probability of successful transmission, during each phase, Node i waits a random time T_w to broadcast its packet, with $\Pr\{T_\mathrm{w} = t\} = 1/\triangle t$ for any $t \in \{1, 2, \ldots, \triangle t\}$. Here $\triangle t$ is an integer and the unit of time is ms. As $\triangle t$ decreases, the nodes will have more difficulties to access the channel. This time constraint due to the MAC layer impacts the performance of the DFD algorithm as will be seen in what follows.

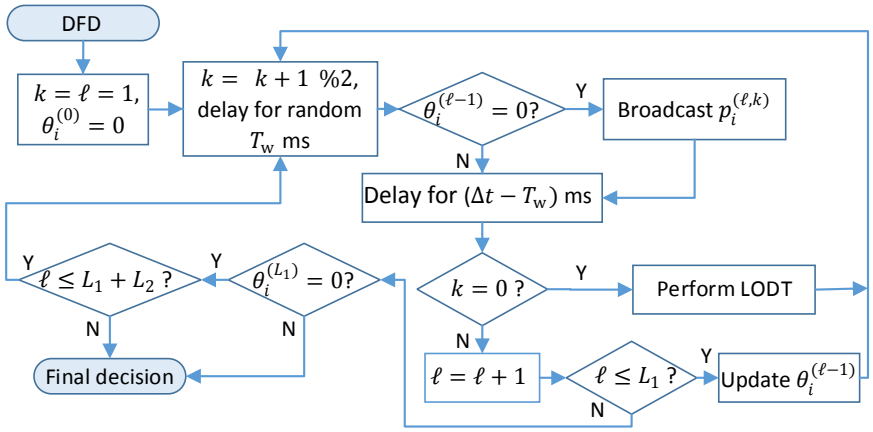

FIGURE 3.15: DFD algorithm performed by Node i

3.6.2 Impact of channel access issues

In a first set of experiments focusing on channel access issues, all the nodes are closely located and can receive packets from all other nodes. However, each node is assigned a random virtual location $\boldsymbol{V}$. Node i and Node j are virtual neighbors if $|\boldsymbol{V}_i - \boldsymbol{V}_j| \leq \bar{r}$, where $\bar{r}$ is the virtual transmission range. $\boldsymbol{V}$ is indicated in the MAC payload. Each node needs to determine whether Node i is its virtual neighbor. If $|\boldsymbol{V}_i - \boldsymbol{V}_j| > \bar{r}$, the packets received from Node i are ignored by Node j. This setting of virtual neighbors, compliant with the assumption in Section , facilities the investigation of the impact of MAC layer. A special node, named *coordinator,* is used to manage the test procedure. At the beginning of each test, the coordinator broadcasts a start message. All the devices start then the DFD algorithm at the same time. After a desired number of independent tests, the coordinator collects the results from the other nodes.

Since all nodes have good sensors in the experiments, some nodes are manually assigned as defective, *i.e.*, their measurements are corrupted by a large constant offset. $n_{\mathrm{s}} = 20$ wireless sensor nodes have been considered, among which $n_{\mathrm{d}} = 3$ nodes are defective. An interesting problem is to compare the performance of algorithm with different number of initial and final rounds as a function of Δt. Independent experiments have been repeated 1000 times for each case. Figure 3.16 presents the experimental results performed on DATASENS platform, as well as the simulation results using Matlab. In the simulations, we set the probability of successfully accessing the channel as $\alpha_1 = \max\{1 - 0.14(n-1), 0\}$, where n denotes the number of nodes that has chosen the same time window t_{W} to access the channel. With this setting, the simulation results and the experimental results are very close in terms of NDR.

For all cases, the NDR decreases as Δt increases, whereas the FAR decreases less significantly. As expected, a large Δt reduces the average number of nodes that have chosen the same t_{W}, and then increases α_1. Therefore the performance of the algorithm becomes better with larger Δt. Moreover, both NDR and FAR become smaller as the number of initial rounds L_1 increases, considering the same Δt.

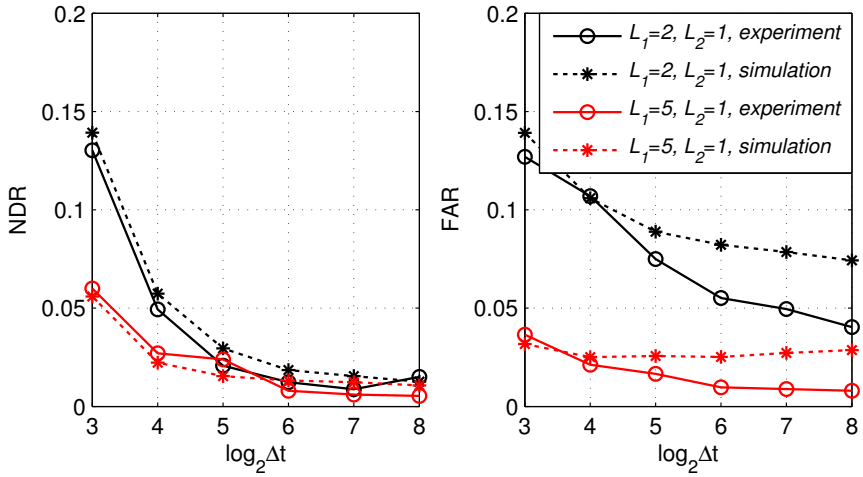

FIGURE 3.16: Average NDR and FAR as functions of Δt for DFD algorithms with different L_1 and L_2.

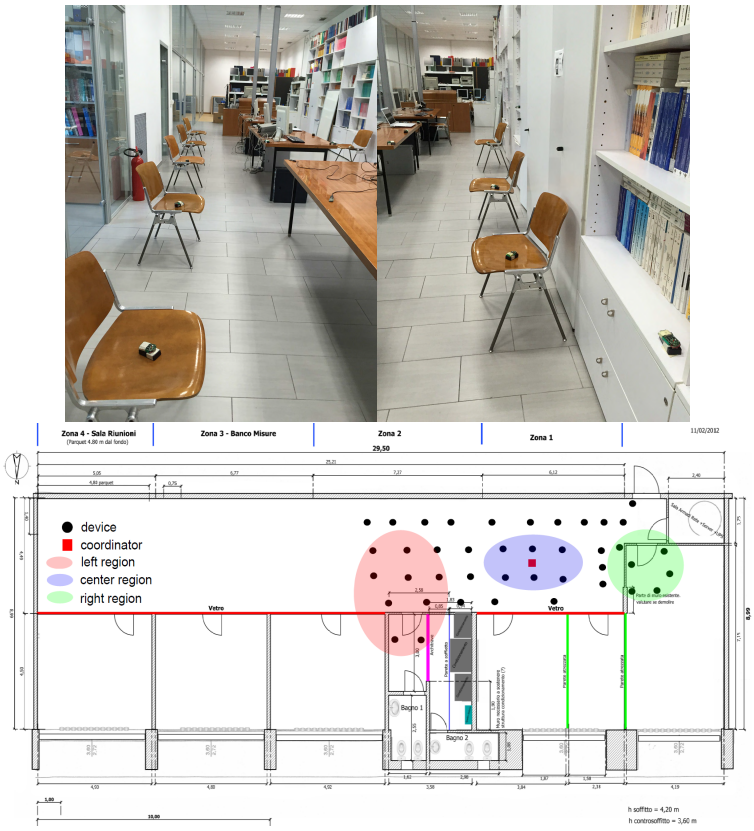

FIGURE 3.17: Node distribution in WiLab

3.6.3 Impact of a real environment

In a second set of experiments, the DFD tests are performed in a more realistic situation where the network is not fully connected. The essential step is to properly choose the transmission power (TxP) of the nodes to have some acceptable coverage distance. The power amplifier (PA) of every device is turned off. The PA of the coordinator keeps functioning and its TxP is set to its maximum to better monitor the test procedure. Figure 3.17 shows the network topology where $n_s = 41$ nodes are randomly deployed over the right side of Wireless Communication Lab (Wilab) of the University of Bologna. The WSN covers an area of $15 \times 4\,\mathrm{m}^2$. The position of nodes remains unchanged, each node has a given probability to be defective in each test. In our tests, the defective probability is set to be 15%, the DFD is performed with $L_1 = 5$ and $L_2 = 1$.

Figure 3.18 illustrates the average DFD performance in different areas and using different TxP, based on 1000 independent tests. The results show that the nodes in the center have lower NDR and FAR than those at sides, considering the same TxP. As is intuitive, the performance of DFD suffers from boundary affects. Three different TxP are considered with their values $P_1 > P_2 > P_3$. Comparing the average NDR and FAR of the nodes in the center, the results highlight that NDR converges faster as TxP decreases, whereas the variation of FAR is not significant. Note that the total execution time of the algorithm is $t_e = 2\Delta t\,(L_1 + L_2)$. The results show that after $t_e = 2 \times 2^6 \times 6 = 768$ms, both NDR and FAR of the nodes in the center are less then 5%, with P_1 and P_2.

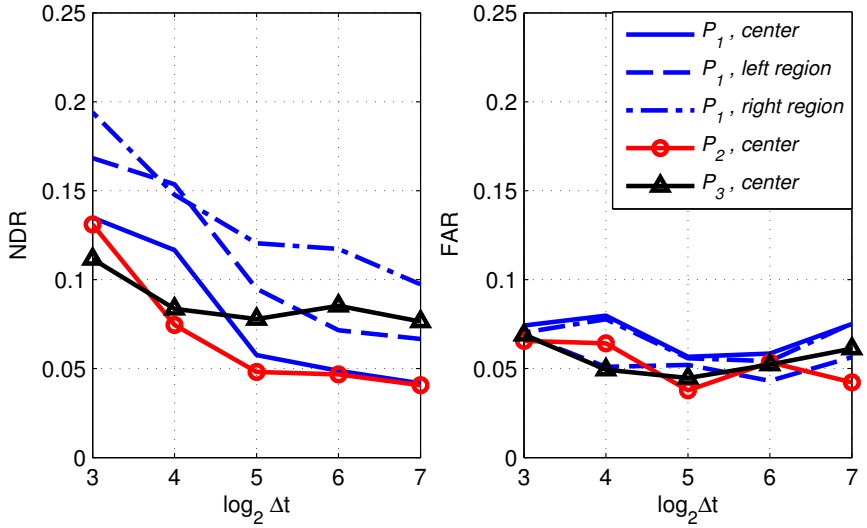

FIGURE 3.18: Average NDR and FAR as functions of $\log_2 \Delta t$ for different transmission power and different area of the testbed.

3.7 Conclusions

This chapter proposed two variants of a two-stage DFD algorithm, which allows each node of a WSN to decide whether its sensor is producing outliers. The performance of the single-decision algorithm has been theoretically characterized. For the iterative variant, sufficient conditions to be satisfied by the LODT are identified to ensure existence of an equilibrium. These results are generic, since no specific form of the LODT needs to be considered. These conditions translate into upper bounds on $\bar{\mu}_{\mathrm{g}}$ and $\bar{\mu}_{\mathrm{d}}$. Conditions for the local asymptotic stability of this equilibrium are also provided. The influence of channel access issues and packet losses has been analyzed.

The DFD algorithm has been tested considering three outliers models and a simple local outlier detection algorithm. When the bounds on $\bar{\mu}_{\mathrm{g}}$ and $\bar{\mu}_{\mathrm{d}}$ are satisfied, the algorithms behave well. When they are not satisfied, some oscillating behavior appears, and an equilibrium is more difficult to reach. Randomly selecting a subset of the received data to perform each LODT addresses this issue.

The performance of both variants have been characterized by simulations, which enables to get some insights on the impact of the algorithm parameters (number of rounds, local test threshold, probability of packet loss) and of the network topology (density of faulty sensors, size of the neighborhood) on the trade-off between P_{D} and P_{FA}. Comparisons with two alternative DFD algorithms have been performed, showing that for comparable or even lower traffic requirements, the proposed algorithm performs better in most of the cases.

The DFD algorithm is also implemented on the real WSN to verify our results. The iterative algorithms with a larger L_1 achieve a better performance under the same time constraint.

A way to account for the effect of node mobility during rounds is, *e.g.*, by an adaptation of ζ. Nevertheless, a precise analysis of the performance of both algorithm in case of significant node mobility is left for future research.

Chapter 4

Distributed Faulty Node Detection in DTNs

Inference and learning over delay tolerant networks (DTNs) is much more complicated than in traditional networks, due their intermittent connectivity. This chapter proposes a fully distributed, easily implementable, and fast convergent approach to allow each DTN node to rapidly identify whether its sensors are producing outliers. Several related work has been discussed in Section 1.2.3.

This chapter is organized as follows. Section 4.1 presents the system model and basic assumptions. Section 4.2 details the DFD algorithm for DTNs. Section 4.3 discusses the transition probabilities between the state values for some reference nodes. Section 4.4 develops the theoretical analysis of the macroscopic evolution of the proportion of nodes in different states. Section 4.5 analyzes the property of the equilibrium obtained from the state equations. Section 4.6 presents the approximation of the proportion of nodes at the equilibrium and discusses the choice of the parameters in the algorithm. Section 4.7 provides some numerical results and Section 4.8 concludes this chapter. Notations are presented in Table 4.1.

4.1 System Model

Consider a set $\mathcal{S}$ of N_S moving nodes equipped with sensors. $\mathcal{D} \subset \mathcal{S}$ represents the subset of nodes with defective sensors producing *outliers*. The *status* of Node i is $\theta_i(t) = 0$ (good node) if all its sensors are good and $\theta_i(t) = 1$ (defective node) if at least one of them is defective. The proportion of nodes with good and defective status are p_0 and p_1, with $p_0 + p_1 = 1$. Each node is not aware of its own status. In what follows, we assume that over the time horizon of the experiment, the status of sensors does not change, *i.e.*, $\theta_i(t) = \theta_i$.

Our aim is (i) to design a distributed algorithm so that each Node i rapidly evaluates an accurate estimate $\widehat{\theta}_i$ of its own status θ_i, as fast as possible, (ii) to provide a theoretical analysis of the behavior of this algorithm.

4.1.1 Communication model

Nodes can exchange information only during the limited time interval in which they are in vicinity. As in [Zhu+10; GLG16; HO+15; Pan+15], we assume that the time interval between two successive meetings follows an exponential distribution with an inter-contact rate λ. Moreover, we assume that

TABLE 4.1: Symbols used in this chapter

θ_i	status of the sensors of Node i
$\widehat{\theta_i}$	estimate of status of the sensors of Node i
λ	inter-contact rate
ν	decision threshold
t	time instant
y_i	outcome of a LODT performed by Node i
q_{D}	detection probability of a LODT
q_{FA}	false alarm probability of a LODT
$c_{\mathrm{m},i}$	number of LODTs performed by Node i
$c_{\mathrm{d},i}$	number of LODTs by Node i resulting in a detection of outliers
M	window size of the LODT results that are considered for the decision
$\mathbf{x}_i$	state of Node i, containing θ_i, $c_{\mathrm{m},i}$, and $c_{\mathrm{d},i}$
$X_\theta^{\ell,k}$	state component of the DTN: proportion of nodes in state $\mathbf{x}_i = (\theta, \ell, k)$ among the nodes with sensors of actual status θ
p_θ	proportion of nodes with sensors of status θ
$p^{\theta\widehat{\theta}}$	proportion of nodes believing its sensors are in status $\widehat{\theta}$, among the nodes with sensors of actual status θ
$\overline{X}_\theta^{\ell,k}$	value of $X_\theta^{\ell,k}$ at equilibrium
$\overline{p}^{\theta\widehat{\theta}}$	value of $p^{\theta\widehat{\theta}}$ at equilibrium
$\widetilde{X}_\theta^{\ell,k}$	approximate value of $X_\theta^{\ell,k}$ at equilibrium
$\widetilde{p}^{\theta\widehat{\theta}}$	approximate value of $p^{\theta\widehat{\theta}}$ at equilibrium

each meeting involves only two nodes. When more than two nodes meet at the same time instant, processing is performed pair-by-pair. These assumptions facilitate the analysis of the proposed DTN-DFD algorithm.

4.1.2 Detection scenario

We assume that during each meeting of a pair of nodes $(i, j) \in \mathcal{S}$, the nodes collect data with their sensors. Each node may or may not transmit its data to the other node. If a node has received data from its neighbor, it may run a LODT (as introduced in Chapter 3) involving its own data and those received from its neighbor. We assume that the spatial and temporal correlation between data is such that only data collected during the meeting of two nodes can be exploited by a LODT. Therefore, previously collected data are not exploited. As a consequence, contrary to the situation of WSN, where n_0 and n_1 may be large, in the DTN scenario, a LODT will involve $n_0, n_1 \in \{0, 1, 2\}$, with $n_0 + n_1 = 2$. In this chapter, one furthermore assumes that $q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(1, 1) \leqslant q_{\mathrm{D}}(0, 2)$, which is reasonable, since the outcome of a LODT is more likely to be 1 as the number of outliers involved increases.

4.2 DTN-DFD algorithm

In the proposed DTN-DFD algorithm, each node manages two counters $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$ initialized at 0 at $t = 0$. Using $c_{\mathrm{m},i}(t)$, Node i counts the number of *meetings* during which it has received

data from its neighbor, and has been able to perform a LODT. Using $c_{\mathrm{d},i}(t)$, it counts the number of LODT tests resulting in the *detection* of outliers. When $c_{\mathrm{d},i}(t)/c_{\mathrm{m},i}(t) \geqslant \nu$, where ν is some constant decision threshold, Node i considers itself as carrying defective sensors, *i.e.*, it sets its own estimate $\widehat{\theta}_i(t) = 1$. Otherwise, it considers that its sensors are good, *i.e.*, $\widehat{\theta}_i(t) = 0$.

When a node with $\widehat{\theta}_i(t) = 1$ encounters another node, it still takes measurements, but it does not send these data to the other node to avoid infecting the network with outliers. Any node, upon receiving data from another node, performs a LODT and updates $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$. As a consequence, a node which meets another node considering itself as defective, transmits its data, but since it does not receive any data, it does not update $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$ at the end of the meeting. Algorithm 3 summarizes the proposed DTN-DFD technique for an arbitrary reference Node i.

The vector $\mathbf{x}_i(t) = (\theta_i, c_{\mathrm{m},i}(t), c_{\mathrm{d},i}(t))$ represents the (microscopic) *state* of each *Node* i. As $t \to \infty$, one has $c_{\mathrm{m},i}(t) \to \infty$, which leads to an infinite number of possible values for $\mathbf{x}_i(t)$ and the global (macroscopic) behavior of the algorithm is difficult to analyze. To limit the number of possible states, one has chosen to consider the evolution of $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$ over a sliding time window containing the time instants of the last M meetings during which Node i has performed a LODT. Algorithm 4 is a modified version of Algorithm 3 accounting for this limited horizon procedure. It involves an additional counter μ indicating the number of LODT performed by Node i. Note that when $\mu < M$, then (4.5) is equivalent to (4.3).

The next sections are devoted to the analysis of Algorithm 4.

4.3 Evolution of the state of a node

The state of Node i is represented by the triple $\mathbf{x}_i(t) = (\theta, c_{\mathrm{m},i}(t), c_{\mathrm{d},i}(t))$. Since $c_{\mathrm{m},i}(t) \in \{0, 1 \dots M\}$ and $c_{\mathrm{d},i}(t) \in \{0 \dots c_{\mathrm{m},i}(t)\}$, the number of values that may be taken by the state of a node is $\sum_{\ell=0}^{M}(\ell+1) = (M+1)(M+2)/2$. The evolution of the state of Node i, conditioned by its status θ_i, follows a Markov model with state transition diagram similar to that shown in Figure 4.1 for $M = 4$.

In particular, there are two chains, one conditioned by $\theta_i = 0$ and the other conditioned by $\theta_i = 1$. Both are characterized by a transient phase for state values with $c_{\mathrm{m},i}(t) < M$, then, a permanent regime starts when $c_{\mathrm{m},i}(t) = M$. With $c_{\mathrm{m},i}(t) = \ell$ and $c_{\mathrm{d},i}(t) = k$, the transitions from State (θ, ℓ, k) to State (θ, ℓ', k') are analyzed in what follows.

Assume that the reference Node i meets a random Node J at time t and define the random event

$$\mathcal{E}_1(t) = \left\{\widehat{\theta}_J(t) = 0\right\}, \tag{4.6}$$

representing the event that the node met believes its status is good. According to (4.4), among the nodes with status θ, the proportion of nodes that believe themselves as good is [1]

$$p^{\theta 0}(t) = X_{\theta}^{0,0}(t) + \sum_{\ell,k,k/\ell<\nu} X_{\theta}^{\ell,k}(t), \tag{4.7}$$

[1] For the sake of simplicity, the dependency of $p^{\theta 0}(t)$ in ν is omitted, as ν is constant during the DTN-DFD algorithm.

Algorithm 3 DTN-DFD algorithm for Node i

1. Initialize at $t_i^0 = 0$, $\widehat{\theta}_i\left(t_i^0\right) = 0$, $c_{\mathrm{m},i}(t_i^0) = c_{\mathrm{d},i}(t_i^0) = 0$, $\kappa = 1$.
2. Do

$$\begin{cases} \widehat{\theta}_i\left(t\right) = \widehat{\theta}_i\left(t_i^{\kappa-1}\right), \\ c_{\mathrm{m},i}\left(t\right) = c_{\mathrm{m},i}\left(t_i^{\kappa-1}\right), \\ c_{\mathrm{d},i}\left(t\right) = c_{\mathrm{d},i}\left(t_i^{\kappa-1}\right), \end{cases} \tag{4.1}$$

$$t = t + \delta t \tag{4.2}$$

   until the κ-th meeting occurs at time $t = t_i^\kappa$ with Node $j^\kappa \in \mathcal{S} \setminus \{i\}$.
3. Perform local measurement of data $m_i\left(t_i^\kappa\right)$.
4. If $\widehat{\theta}_i\left(t_i^\kappa\right) = 0$, then transmit $m_i\left(t_i^\kappa\right)$ to Node j^κ.
5. If data m_{j^κ} have been received from Node j^κ, then
   (a) Perform a LODT with outcome $y_i\left(t_i^\kappa\right)$.
   (b) Update $c_{\mathrm{m},i}$ and $c_{\mathrm{d},i}$ according to

$$\begin{cases} c_{\mathrm{m},i}(t_i^\kappa) = c_{\mathrm{m},i}(t_i^{\kappa-1}) + 1 \\ c_{\mathrm{d},i}(t_i^\kappa) = c_{\mathrm{d},i}(t_i^{\kappa-1}) + y_i\left(t_i^\kappa\right) \end{cases} \tag{4.3}$$

   (c) Update $\widehat{\theta}_i$ as follows

$$\widehat{\theta}_i(t_i^\kappa) = \begin{cases} 1 & \text{if } c_{\mathrm{d},i}(t_i^\kappa)/c_{\mathrm{m},i}(t_i^\kappa) \geqslant \nu, \\ 0 & \text{else.} \end{cases} \tag{4.4}$$

6. $\kappa = \kappa + 1$.
7. Go to 2.

Algorithm 4 Sliding-Window DTN-DFD algorithm for Node i

1. Initialize $t_i^0 = 0$, $\widehat{\theta}_i\left(t_i^0\right) = 0$, $c_{\mathrm{m},i}(t_i^0) = c_{\mathrm{d},i}(t_i^0) = 0$, $\kappa = 1$, and $\mu = 0$.
2. Do (4.1)-(4.2) until the κ-th meeting occurs at time t_i^κ with Node $j^\kappa \in \mathcal{S} \setminus \{i\}$.
3. Perform local measurement of data $m_i\left(t_i^\kappa\right)$.
4. If $\widehat{\theta}_i\left(t_i^\kappa\right) = 0$, then transmit $m_i\left(t_i^\kappa\right)$ to Node j^κ.
5. If data m_{j^κ} have been received from Node j^κ, then
   (a) $\mu = \mu + 1$. Perform a LODT with outcome y_i^μ.
   (b) Update $c_{\mathrm{m},i}$ and $c_{\mathrm{d},i}$ as
   $$\begin{cases} c_{\mathrm{m},i}(t_i^\kappa) = \min\{\mu, M\}, \\ c_{\mathrm{d},i}(t_i^\kappa) = \sum_{m=\max\{1,\mu-M+1\}}^{\mu} y_i^m. \end{cases} \tag{4.5}$$
   (c) Update $\widehat{\theta}_i$ using (4.4).
6. $\kappa = \kappa + 1$.
7. Go to 2.

where $p^{10}(t)$ is in fact the *non-detection rate* (NDR) of the nodes with defective sensors at time t. Assuming that the nodes are randomly spread, the probability that Node J believes it has only good sensors conditioned to its true status is

$$\mathbb{P}\left(\widehat{\theta}_J(t) = 0 | \theta_J(t) = \theta\right) = p^{\theta 0}(t), \tag{4.8}$$

and then

$$\mathbb{P}\{\mathcal{E}_1(t)\} = p_0 p^{00}(t) + p_1 p^{10}(t). \tag{4.9}$$

Similarly, introducing $\overline{\mathcal{E}_1}(t) = \{\widehat{\theta}_J(t) = 1\}$, among the nodes with sensors in status θ, the proportion of nodes with $\widehat{\theta}_j(t) = 1$ is

$$p^{\theta 1}(t) = \sum_{\ell,k,k/\ell \geqslant \nu} X_\theta^{\ell,k}(t), \tag{4.10}$$

where $p^{01}(t)$ and $p^{11}(t)$ represent the *false alarm rate* (FAR) and the *detection rate* (DR) respectively. From (4.10), one gets

$$\mathbb{P}\left\{\overline{\mathcal{E}_1}(t)\right\} = p_0 p^{01}(t) + p_1 p^{11}(t). \tag{4.11}$$

Since Node i performs an LODT only when it meets a node J with $\widehat{\theta}_J(t) = 0$, one introduces the random event

$$\mathcal{E}_2^\theta(t) = \left\{Y_i(t) = 1 \mid \theta_i = \theta, \widehat{\theta}_J(t) = 0\right\}, \tag{4.12}$$

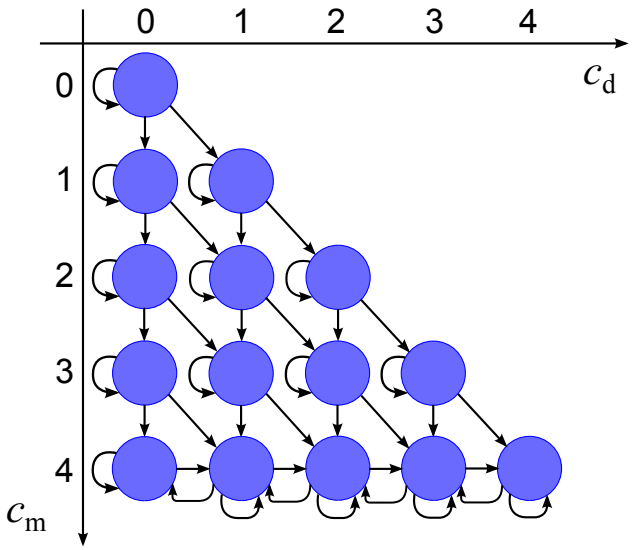

FIGURE 4.1: Example of Markov model for the evolution of the state of a node when $M = 4$.

for the reference node with actual status θ. The statistical properties of the outcome $Y_i(t)$ of the LODT depend only on θ_i and θ_j. For example, when Node i has good sensors, one has

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_2^0(t)\right\} &= \sum_{\varphi=0}^{1} \mathbb{P}\left\{Y_i(t)=1, \theta_J=\varphi \mid \theta_i=0, \widehat{\theta}_J(t)=0\right\} \\ &\overset{(a)}{=} \sum_{\varphi=0}^{1} \mathbb{P}\left\{Y_i(t)=1 \mid \theta_i=0, \theta_J=\varphi\right\} \mathbb{P}\left\{\theta_J=\varphi \mid \widehat{\theta}_J(t)=0\right\} \\ &\overset{(b)}{=} \frac{p_0 q_{\mathrm{FA}}(2)\, p^{00}(t) + p_1 q_{\mathrm{D}}(1,1)\, p^{10}(t)}{p_0 p^{00}(t) + p_1 p^{10}(t)}. \end{aligned} \tag{4.13}$$

In (4.13-a), one uses the fact that the LODT outcome is not influenced by the estimate of the status of a node and that in $\mathbb{P}\left\{\theta_J=\varphi \mid \theta_i=0, \widehat{\theta}_J(t)=0\right\}$, the status of Node J, does not depend on θ_i. In (4.13-b),

$$\begin{cases} \mathbb{P}\left\{Y_i(t)=1 \mid \theta_i=0, \theta_J=0\right\} = q_{\mathrm{FA}}(2), \\ \mathbb{P}\left\{Y_i(t)=1 \mid \theta_i=0, \theta_J=1\right\} = q_{\mathrm{D}}(1,1). \end{cases} \tag{4.14}$$

Moreover,

$$\mathbb{P}\left\{\theta_J=\varphi \mid \widehat{\theta}_J(t)=0\right\} = \frac{\mathbb{P}\left\{\widehat{\theta}_J(t)=0 \mid \theta_J=\varphi\right\}\mathbb{P}\{\theta_J=\varphi\}}{\sum_{\phi=0}^{1}\mathbb{P}\left\{\widehat{\theta}_J(t)=0 \mid \theta_J=\phi\right\}\mathbb{P}\{\theta_J=\phi\}} = \frac{p_\varphi p^{\varphi 0}(t)}{p_0 p^{00}(t) + p_1 p^{10}(t)}.$$

Now, if Node i has defective sensors, (4.12) can be expressed as

$$\mathbb{P}\left\{\mathcal{E}_2^1(t)\right\} = \frac{p_0 q_{\mathrm{D}}(1,1)\, p^{00}(t) + p_1 q_{\mathrm{D}}(0,2)\, p^{10}(t)}{p_0 p^{00}(t) + p_1 p^{10}(t)}. \tag{4.15}$$

Similarly, one may introduce the random event

$$\mathcal{E}_3^{\theta}(t) = \left\{ Y_i(t) = 0 \mid \theta_i = \theta, \widehat{\theta}_J(t) = 0 \right\}, \tag{4.16}$$

and show that

$$\mathbb{P}\left\{\mathcal{E}_3^{\theta}(t)\right\} = \begin{cases} \frac{p_0(1-q_{\mathrm{FA}}(2))p^{00}(t)+p_1(1-q_{\mathrm{D}}(1,1))p^{10}(t)}{p_0p^{00}(t)+p_1p^{10}(t)}, & \text{if } \theta = 0, \\ \frac{p_0(1-q_{\mathrm{D}}(1,1))p^{00}(t)+p_1(1-q_{\mathrm{D}}(0,2))p^{10}(t)}{p_0p^{00}(t)+p_1p^{10}(t)}, & \text{if } \theta = 1. \end{cases} \tag{4.17}$$

Define $\pi_{\theta}^{\delta_{\mathrm{m}},\delta_{\mathrm{d}}}$ as the transition probability from State (θ, ℓ, k) to State $(\theta, \ell + \delta_{\mathrm{m}}, k + \delta_{\mathrm{d}})$, where $\theta \in \{0, 1\}$. One has $\delta_{\mathrm{m}} \in \{0, 1\}$ since ℓ may either increase ($\delta_{\mathrm{m}} = 1$) in the transient regime or remain constant ($\delta_{\mathrm{m}} = 0$) in the permanent regime. One has $\delta_{\mathrm{d}} \in \{-1, 0, 1\}$, depending on the value of the last LODT outcome and on the value of the $M+1$-th last LODT outcome, which is no more considered in the permanent regime. Thus, $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) \in \{(0,0), (0,1), (0,-1)(1,0), (1,1), (1,-1)\}$. Note that $\pi_{\theta}^{\delta_{\mathrm{m}},\delta_{\mathrm{d}}}$ depends on the current state of the reference node, but also on the current proportion of active (good and defective) nodes. Therefore, the transition probabilities are denoted as $\pi_{\theta}^{\delta_{\mathrm{m}},\delta_{\mathrm{d}}}(t, \ell, k)$, where t is the time instant, $c_{\mathrm{m},i}(t) = \ell$, and $c_{\mathrm{d},i}(t) = k$. Depending on the value of ℓ, two different cases are considered in Section 4.3.1 and in Section 4.3.2, respectively corresponding to the transient and permanent regimes.

4.3.1 Case I, $\ell < M$

In the transient regime, when $c_{\mathrm{m},i}(t) = \ell < M$, $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$ are updated according to (4.3) whenever Node J with $\widehat{\theta}_J(t) = 0$ is met,. The only possibility that leads to $\delta_{\mathrm{m}} = 0$ is the event $\overline{\mathcal{E}_1}$, *i.e.*, Node i meets Node J with $\widehat{\theta}_J(t) = 1$. As a consequence, no LODT is performed by Node i. Therefore, for any $\theta \in \{0, 1\}$,

$$\pi_{\theta}^{0,0}(t, \ell, k) = \mathbb{P}\left\{\overline{\mathcal{E}_1}(t)\right\} = p_0 p^{01}(t) + p_1 p^{11}(t), \tag{4.18}$$

where $p^{\theta 1}(t)$ is defined by (4.10).

A state transition occurs with $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) = (1, 1)$ when Node i with status $\theta_i = \theta$ meets Node J with $\widehat{\theta}_J(t) = 0$ *and* when the LODT yields $y_i(t) = 1$. Since the two events are independent, one has

$$\pi_{\theta}^{1,1}(t, \ell, k) = \mathbb{P}\left\{Y_i(t) = 1, \widehat{\theta}_J(t) = 0 | \theta_i = \theta\right\} = \mathbb{P}\left\{\mathcal{E}_1(t)\right\} \mathbb{P}\left\{\mathcal{E}_2^{\theta}(t)\right\}. \tag{4.19}$$

Depending on the value of θ_i, using (4.9), (4.13), and (4.15), one may rewrite (4.19) as

$$\pi_{\theta}^{1,1}(t, \ell, k) = \begin{cases} p_0 q_{\mathrm{FA}}(2) p^{00}(t) + p_1 q_{\mathrm{D}}(1,1) p^{10}(t), & \text{if } \theta = 0, \\ p_0 q_{\mathrm{D}}(1,1) p^{00}(t) + p_1 q_{\mathrm{D}}(0,2) p^{10}(t), & \text{if } \theta = 1. \end{cases} \tag{4.20}$$

Finally, $\pi_\theta^{1,0}(t,\ell,k) = \mathbb{P}\left\{Y_i(t)=0, \widehat{\theta}_J(t)=0 | \theta_i=\theta\right\}$ is obtained similarly from (4.17)

$$\pi_\theta^{1,0}(t,\ell,k) = \begin{cases} p_0\left(1-q_{\mathrm{FA}}(2)\right)p^{00}(t) + p_1\left(1-q_{\mathrm{D}}(1,1)\right)p^{10}(t), & \text{if } \theta=0, \\ p_0\left(1-q_{\mathrm{D}}(1,1)\right)p^{00}(t) + p_1\left(1-q_{\mathrm{D}}(0,2)\right)p^{10}(t), & \text{if } \theta=1. \end{cases} \tag{4.21}$$

4.3.2 Case II, $c_{\mathbf{m},i}(t) = M$

In the permanent regime, $c_{\mathrm{m},i}(t) = M$ and does not increase any more, thus $\delta_{\mathrm{m}} = 0$. In Algorithm 4, μ is the number of LODTs performed by Node i up to time t. When $\mu \geqslant M$, only the last M LODT outcomes are considered: LODT outcomes y_i^m with $m \leqslant \mu - M$ are no more considered.

To determine the value taken by $\delta_{\mathrm{d}} \in \{-1,0,1\}$ after the μ-th LODT, consider the random event

$$\mathcal{E}_4^1(t) = \left\{ Y_i^{\mu-M} = 1 \mid \sum_{m=\mu-M}^{\mu-1} Y_i^m = k \right\}, \tag{4.22}$$

which corresponds to a situation where one knows that k LODTs where positive among the last M tests and the LODT that will be ignored, once the new LODT outcome is available, also concluded in the presence of defective sensors. $\mathbb{P}\left\{\mathcal{E}_4^1(t)\right\}$ is relatively complex to evaluate, since $\mathbb{P}\{Y_i^n = 1\}$ is time-varying according to (4.13-4.15). In what follows, we assume that LODT outcomes with $Y_i^m = 1$ are independently distributed over the time horizon corresponding to $m = \mu-M, \ldots, \mu-1$. One obtains then

$$\mathbb{P}\left\{\mathcal{E}_4^1(t)\right\} = \frac{k}{M}. \tag{4.23}$$

This approximation is exact in steady-state, when the $X_\theta^{\ell,k}$s do not vary any more. Similarly, define $\mathcal{E}_4^0(t) = \left\{Y_i^{\mu-M} = 0 \mid \sum_{m=\mu-M}^{\mu-1} Y_i^m = k\right\}$. Making the same assumption used to get (4.23), one has

$$\mathbb{P}\left\{\mathcal{E}_4^0(t)\right\} = 1 - \mathbb{P}\left\{\mathcal{E}_4^1(t)\right\} \approx \frac{M-k}{M}. \tag{4.24}$$

Assume that the $(\mu-M)$-th LODT performed by Node i occurred at time $\tilde{t}$, then $y_i^{\mu-M}$ can also be denoted as $y_i\left(\tilde{t}\right)$ and the transition related to $c_{\mathrm{d},i}$ is such that $\delta_{\mathrm{d}} = y_i(t) - y_i\left(\tilde{t}\right) \in \{-1,0,1\}$.

To have $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) = (0,1)$, three independent events have to occur: 1) the encountered Node J believes it is good at time t, *i.e.*, $\mathcal{E}_1(t)$; 2) $y_i(t) = 1$, *i.e.*, $\mathcal{E}_2(t)$; 3) $y_i\left(\tilde{t}\right) = 0$, *i.e.*, $\mathcal{E}_4^0(t)$. Thus the transition probability may be expressed as

$$\pi_\theta^{0,1}(t,M,k) = \mathbb{P}\left\{\mathcal{E}_1(t)\right\}\mathbb{P}\left\{\mathcal{E}_2^\theta(t)\right\}\mathbb{P}\left\{\mathcal{E}_4^0(t)\right\}. \tag{4.25}$$

Using (4.9), (4.13), (4.15), and (4.23) in (4.25), one gets

$$\pi_\theta^{0,1}(t,M,k) = \begin{cases} \left(p_0 q_{\mathrm{FA}}(2)\, p^{00}(t) + p_1 q_{\mathrm{D}}(1,1)\, p^{10}(t)\right)\frac{M-k}{M}, & \text{if } \theta=0, \\ \left(p_0 q_{\mathrm{D}}(1,1)\, p^{00}(t) + p_1 q_{\mathrm{D}}(0,2)\, p^{10}(t)\right)\frac{M-k}{M}, & \text{if } \theta=1. \end{cases} \tag{4.26}$$

Consider now $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) = (0,-1)$. To have such transition, the three following independent events should occur: 1) $\mathcal{E}_1(t)$; 2) $y_i(t) = 0$, *i.e.*, $\mathcal{E}_3(t)$; 3) $y_i\left(\tilde{t}\right) = 1$, *i.e.*, $\mathcal{E}_4^1(t)$. Thus, the

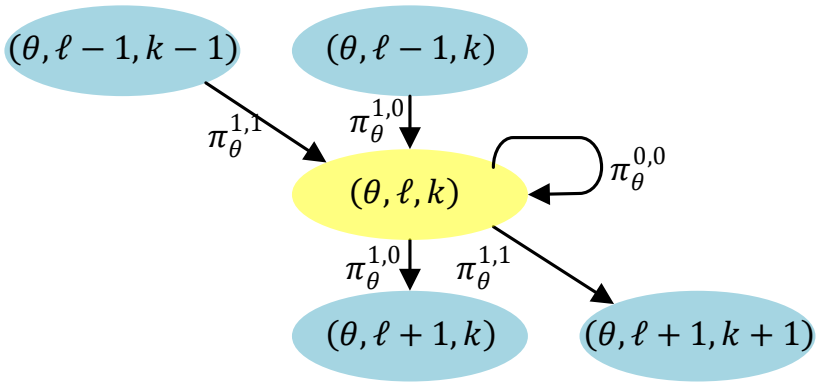

FIGURE 4.2: Transient regime: Possible state transitions from and to state (θ, ℓ, k) when $0 < \ell < M$ and $0 < k < \ell$

transition probability is

$$\begin{aligned}\pi_\theta^{0,-1}(t, M, k) &= \mathbb{P}\{\mathcal{E}_1(t)\}\,\mathbb{P}\left\{\mathcal{E}_3^\theta(t)\right\}\mathbb{P}\left\{\mathcal{E}_4^1(t)\right\} \\ &= \begin{cases} \left(p_0(1-q_{\mathrm{FA}}(2))\,p^{00}(t) + p_1(1-q_{\mathrm{D}}(1,1))\,p^{10}(t)\right)\frac{k}{M}, & \text{if } \theta = 0, \\ \left(p_0(1-q_{\mathrm{D}}(1,1))\,p^{00}(t) + p_1(1-q_{\mathrm{D}}(0,2))\,p^{10}(t)\right)\frac{k}{M}, & \text{if } \theta = 1. \end{cases}\end{aligned} \tag{4.27}$$

Considering the last transition $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) = (0,0)$. To obtain the expression of $\pi_\theta^{0,0}(t, M, k)$, one needs to introduce (4.26-4.27) into the following

$$\pi_\theta^{0,0}(t, M, k) = 1 - \pi_\theta^{0,1}(t, M, k) - \pi_\theta^{0,-1}(t, M, k)\,. \tag{4.28}$$

4.4 Macroscopic evolution of the DTN state

At time t, among the nodes with status $\theta \in \{0,1\}$, let $X_\theta^{\ell,k}(t)$ be the proportion of nodes in state (θ, ℓ, k). All node state transition probabilities evaluated in Section 4.3 are now used to determine the evolution of the *DTN state* components, *i.e.*, of the various proportions of nodes $X_0^{\ell,k}(t)$ and $X_1^{\ell,k}(t)$ in the corresponding states, with $\ell = 0, \ldots, M$ and $k \leqslant \ell$.

Considering an inter-contact rate λ and a well-mixed population of nodes, during a short time interval $[t, t+\delta t]$ the number of nodes with state (θ, ℓ, k) that will meet another node is $\lambda p_\theta N_S X_\theta^{\ell,k}(t)\delta t$.

When $0 < \ell < k < M$, these nodes will switch to the states $(\theta, \ell + \delta_{\mathrm{m}}, k + \delta_{\mathrm{d}})$, with $(\delta_{\mathrm{m}}, \delta_{\mathrm{d}}) \in \{(0,0), (1,0), (1,1)\}$ with a probability $\pi_\theta^{\delta_{\mathrm{m}},\delta_{\mathrm{d}}}(t, \ell, k)$. Moreover, nodes in the states $(\theta, \ell-1, k-1)$ and $(\theta, \ell-1, k)$ that have met an other node in the time interval $[t, t+\delta t]$ may reach state (θ, ℓ, k), respectively with a probability $\pi_\theta^{1,1}(t, \ell-1, k-1)$ and $\pi_\theta^{1,0}(t, \ell-1, k)$, see Figure 4.2.

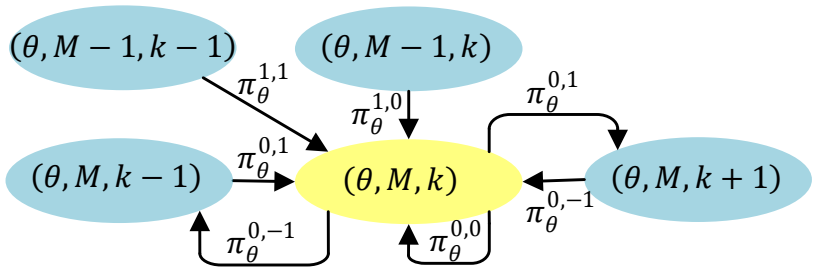

FIGURE 4.3: Permanent regime: Possible state transitions from and to State (θ, M, k) when $0 < k < M$

As a consequence, at time $t + \delta t$, the number of nodes in State (θ, ℓ, k) may be expressed as follows

$$\begin{aligned} p_\theta N_S X_\theta^{\ell,k}(t+\delta t) = p_\theta N_S X_\theta^{\ell,k}(t) + \lambda \delta t p_\theta N_S \Big(-X_\theta^{\ell,k}(t) \left(\pi_\theta^{1,0}(t,\ell,k) + \pi_\theta^{1,1}(t,\ell,k) \right) \\ + X_\theta^{\ell-1,k-1}(t) \pi_\theta^{1,1}(t,\ell-1,k-1) + X_\theta^{\ell-1,k}(t) \pi_\theta^{1,0}(t,\ell-1,k) \Big). \end{aligned} \tag{4.29}$$

The evolution of $X_\theta^{\ell,k}(t)$ is then described by the following differential equation, where the time dependency is omitted to lighten notations

$$\begin{aligned} \frac{dX_\theta^{\ell,k}}{dt} = -\lambda X_\theta^{\ell,k} \left(\pi_\theta^{1,0}(\ell,k) + \pi_\theta^{1,1}(\ell,k) \right) \\ + \lambda X_\theta^{\ell-1,k-1} \pi_\theta^{1,1}(\ell-1,k-1) + \lambda X_\theta^{\ell-1,k} \pi_\theta^{1,0}(\ell-1,k). \end{aligned} \tag{4.30}$$

When $\ell = M$ and $0 < k < M$, nodes in state (θ, M, k) will switch to the states $(\theta, M, k + \delta_{\mathrm{d}})$, $\delta_{\mathrm{d}} \in \{-1, 0, 1\}$ with a probability $\pi_\theta^{0,\delta_{\mathrm{d}}}(t, M, k)$. Nodes in the states $(\theta, M-1, k-1)$ and $(\theta, M-1, k)$ that have met an other node in the time interval $[t, t+\delta t]$ may reach state (θ, M, k), respectively with a probability $\pi_\theta^{1,1}(t, M-1, k-1)$ and $\pi_\theta^{1,0}(t, M-1, k)$, see Figure 4.3. As a consequence, the evolution of $X_\theta^{M,k}(t)$ can be described by

$$\begin{aligned} \frac{dX_\theta^{M,k}}{dt} = -\lambda X_\theta^{M,k} \left(\pi_\theta^{0,1}(M,k) + \pi_\theta^{0,-1}(M,k) \right) + \lambda X_\theta^{M-1,k-1} \pi_\theta^{1,1}(M-1,k-1) \\ + \lambda X_\theta^{M-1,k} \pi_\theta^{1,0}(M-1,k) + \lambda X_\theta^{M,k-1} \pi_\theta^{0,1}(M,k-1) + \lambda X_\theta^{M,k+1} \pi_\theta^{0,-1}(M,k+1). \end{aligned} \tag{4.31}$$

Similar derivations can be made for the remaining DTN state components to obtain

$$\begin{cases} \frac{dX_\theta^{0,0}}{dt} \overset{(a)}{=} -\lambda X_\theta^{0,0} \left(\pi_\theta^{1,0}(0,0) + \pi_\theta^{1,1}(0,0) \right), \\ \frac{dX_\theta^{\ell,0}}{dt} \overset{(b)}{=} \lambda \left(-X_\theta^{\ell,0} \left(\pi_\theta^{1,0}(\ell,0) + \pi_\theta^{1,1}(\ell,0) \right) + X_\theta^{\ell-1,0} \pi_\theta^{1,0}(\ell-1,0) \right) \\ \frac{dX_\theta^{\ell,\ell}}{dt} \overset{(c)}{=} \lambda \left(-X_\theta^{\ell,\ell} \left(\pi_\theta^{1,0}(\ell,\ell) + \pi_\theta^{1,1}(\ell,\ell) \right) + X_\theta^{\ell-1,\ell-1} \pi_\theta^{1,1}(\ell-1,\ell-1) \right), \\ \frac{dX_\theta^{M,0}}{dt} \overset{(d)}{=} \lambda \left(-X_\theta^{M,0} \pi_\theta^{0,1}(M,0) + X_\theta^{M-1,0} \pi_\theta^{1,0}(M-1,0) + X_\theta^{M,1} \pi_\theta^{0,-1}(M,1) \right), \\ \frac{dX_\theta^{M,M}}{dt} \overset{(e)}{=} \lambda \left(-X_\theta^{M,M} \pi_\theta^{0,-1}(M,M) + X_\theta^{M,M-1} \pi_\theta^{0,1}(M,M-1) + X_\theta^{M-1,M-1} \pi_\theta^{1,1}(M-1,M-1) \right), \end{cases} \tag{4.32}$$

for any $\ell = 1 \ldots M-1$, with the initial conditions $X_\theta^{0,0}(0) = 1$ and $X_\theta^{\ell,k}(0) = 0$, $\forall \ell, k \neq 0$.

The state equation (4.32) is nonlinear, since each $\pi_\theta^{\delta\mathrm{m},\delta\mathrm{d}}$ depends on $X_\theta^{\ell,k}$, see (4.7) and (4.10).

4.5 Analysis of the DTN state equations

In what follows, the asymptotic behavior of the DTN state equations (4.32) is characterized. Algorithm 4 may drive $X_\theta^{\ell,k}$ to an equilibrium $\overline{X}_\theta^{\ell,k}$ at which the proportions of nodes in different states $X_\theta^{\ell,k}(t)$ do not vary any more. As a consequence, $p^{\theta 0}(t)$ defined in (4.7) also tends to an equilibrium $\overline{p}^{\theta 0}$.

4.5.1 Equilibrium of $X_\theta^{\ell,k}$

One investigates first the evolution of $X_\theta^{\ell,k}(t)$ when $\ell < M$. As shown in the following proposition, the DTN state always reaches the permanent regime.

Proposition 4.1. *For any $\ell < M$ and $k \leqslant \ell$, $\lim_{t\to\infty} X_\theta^{\ell,k}(t) = 0$.*

Proof. See Appendix C.1. □

From Proposition 4.1, the only possible value at equilibrium of $X_\theta^{\ell,k}(t)$ when $\ell < M$ is 0. Thus $\overline{p}^{\theta 0}$ may be written as

$$\overline{p}^{\theta 0} = \sum_{k:k/M<\nu} \overline{X}_\theta^{M,k}. \tag{4.33}$$

Denote $\overline{\mathbf{p}} = \left(\overline{p}^{00}, \overline{p}^{10}\right) \in \mathcal{P}_0$ with

$$\mathcal{P}_0 = \{(x,y) \in [0,1]\times[0,1] \text{ and } (x,y) \neq (0,0)\} \tag{4.34}$$

and consider the functions

$$h_0(\overline{\mathbf{p}}) = \frac{p_0 q_{\mathrm{FA}}(2)\,\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\,\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}, \tag{4.35}$$

$$h_1(\overline{\mathbf{p}}) = \frac{p_0 q_{\mathrm{D}}(1,1)\,\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\,\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}, \tag{4.36}$$

$$F_\theta(\overline{\mathbf{p}}) = \sum_{k=0}^{\lceil M\nu\rceil-1} \binom{M}{k} \left(h_\theta(\overline{\mathbf{p}})\right)^k \left(1-h_\theta(\overline{\mathbf{p}})\right)^{M-k}, \tag{4.37}$$

and $\mathbf{F}(\overline{\mathbf{p}}) = (F_0(\overline{\mathbf{p}}), F_1(\overline{\mathbf{p}}))$. The following proposition provides a non-linear equation that has to be satisfied by $\overline{\mathbf{p}}$. The various $\overline{X}_\theta^{M,d}$ at equilibrium are easily deduced from the solutions of the mentioned equation.

Proposition 4.2. *Assume that the dynamic system described by (4.30-4.32) admits some equilibrium $\overline{X}_\theta^{\ell,d}$, then $\overline{\mathbf{p}} \in \mathcal{P}_0$ is the solution of*

$$\overline{\mathbf{p}} = \mathbf{F}(\overline{\mathbf{p}}), \tag{4.38}$$

and for any $\theta \in \{0,1\}$ *and* $k \leqslant \ell$,

$$\overline{X}_{\theta}^{\ell,k} = \begin{cases} 0, & \forall \ell < M, \\ \binom{M}{k} \left(h_\theta\left(\overline{\mathbf{p}}\right)\right)^k \left(1 - h_\theta\left(\overline{\mathbf{p}}\right)\right)^{M-k}, & \ell = M. \end{cases} \tag{4.39}$$

Proof. See Appendix C.2. □

4.5.2 Existence and unicity of the equilibrium point

Now we investigate the existence and the unicity of the solution of (4.38), which is rewritten in detail as the following equations

$$\begin{cases} \overline{p}^{00} = F_0\left(\overline{p}^{00}, \overline{p}^{10}\right) = \sum\limits_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\mathrm{FA}}(2)\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^k \left(\frac{p_0(1-q_{\mathrm{FA}}(2))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(1,1))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^{M-k}, \\ \overline{p}^{10} = F_0\left(\overline{p}^{00}, \overline{p}^{10}\right) = \sum\limits_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\mathrm{D}}(1,1)\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^k \left(\frac{p_0(1-q_{\mathrm{D}}(1,1))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(0,2))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^{M-k}. \end{cases} \tag{4.40}$$

For that purpose, using fixed-point theorems, one may alternatively show that for all $\mathbf{p}(0) = \left(p^{00}(0), p^{10}(0)\right) \in \mathcal{P}_0$, the discrete-time system

$$\begin{cases} p^{00}(n+1) = F_0\left(p^{00}(n), p^{10}(n)\right), \\ p^{10}(n+1) = F_1\left(p^{00}(n), p^{10}(n)\right). \end{cases} \tag{4.41}$$

converges to a unique equilibrium point $\left(\overline{p}^{00}, \overline{p}^{10}\right)$, which is then solution of (4.40).

One first shows the existence of an equilibrium using Brouwer's fixed-point theorem [GD13] in the following proposition.

Proposition 4.3. *For any* $\nu \in [0,1]$, (4.40) *always admits a solution, which is an equilibrium point of the dynamical system* (4.30)-(4.32).

Before proving Proposition 4.3, one first shows that $p^{00}(n)$ and $p^{10}(n)$ are contained in intervals with lower and upper bounds increasing (resp. decreasing) with n.

Lemma 4.1. *For any* $n \in \mathbb{N}^*$ *and* $\theta \in \{0,1\}$, *one has*

$$p_{\min}^{\theta 0}(n) \leqslant p^{\theta 0}(n) \leqslant p_{\max}^{\theta 0}(n),$$

with $p_{\min}^{\theta 0}(0) = 0$, $p_{\max}^{\theta 0}(0) = 1$, *and*

$$\begin{cases} p_{\min}^{\theta 0}(n+1) = F_\theta\left(p_{\min}^{00}(n), p_{\max}^{10}(n)\right), & \forall n \in \mathbb{N}^+, \\ p_{\max}^{\theta 0}(n+1) = F_\theta\left(p_{\max}^{00}(n), p_{\min}^{10}(n)\right), & \forall n \in \mathbb{N}^+. \end{cases} \tag{4.42}$$

Moreover,

$$p_{\min}^{00}(n+1) > p_{\min}^{00}(n), \qquad p_{\max}^{00}(n+1) < p_{\max}^{00}(n). \tag{4.43}$$

Proof. See Appendix C.3. □

Using Lemma 4.1, one can now prove Proposition 4.3.

Proof. F_0 and F_1 are both continuous functions. For some $n > 0$, consider the set $\mathcal{P}_n = \left[p^{00}_{\min}(n), p^{00}_{\max}(n)\right] \times \left[p^{10}_{\min}(n), p^{10}_{\max}(n)\right]$, where $p^{\theta 0}_{\min}(n)$ and $p^{\theta 0}_{\max}(n)$ are defined in (4.42). For any $\mathbf{p} = \left(p^{00}, p^{10}\right) \in \mathcal{P}_n$, one can prove using Lemma 4.1 that $\mathbf{F}(\mathbf{p}) \in \mathcal{P}_n$. Thus $\mathbf{F}$ maps $\mathcal{P}_n$ to $\mathcal{P}_n$. Applying Brouwer's fixed-point theorem, $\mathbf{F}$ admits a fixed point and Proposition 4.3 is proved. □

Sufficient conditions on p_0, p_1, q_{D}, q_{FA}, M and ν are then provided to ensure the uniqueness of this equilibrium by applying Banach's fixed-point theorem [Ban22].

Proposition 4.4. *If there exists some N', such that $\forall \theta \in \{0,1\}$ and $\forall n > N'$, one has*

$$c_\theta(q_{\mathrm{FA}}(2), q_{\mathrm{D}}(0,2), q_{\mathrm{D}}(1,1), p_1, M, \nu, n) < 1, \tag{4.44}$$

where c_0 and c_1 are defined as

$$\begin{aligned} &c_0(q_{\mathrm{FA}}(2), q_{\mathrm{D}}(0,2), q_{\mathrm{D}}(1,1), p_1, M, \nu, n) \\ &= \frac{M\left(q_{\mathrm{D}}(1,1) - q_{\mathrm{FA}}(2)\right) p_0 p_1 p^{00}_{\max}(n)\, p^{10}_{\max}(n)}{\left(p_0 p^{00}_{\min}(n) + p_1 p^{10}_{\min}(n)\right)\left((1 - q_{\mathrm{FA}}(2)) p_0 p^{00}_{\min}(n) + (1 - q_{\mathrm{D}}(1,1)) p_1 p^{10}_{\min}(n)\right)}, \end{aligned} \tag{4.45}$$

$$\begin{aligned} &c_1(q_{\mathrm{FA}}(2), q_{\mathrm{D}}(0,2), q_{\mathrm{D}}(1,1), p_1, M, \nu, n) \\ &= \frac{M\left(q_{\mathrm{D}}(0,2) - q_{\mathrm{D}}(1,1)\right) p_0 p_1 p^{00}_{\max}(n)\, p^{10}_{\max}(n)}{\left(p_0 p^{00}_{\min}(n) + p_1 p^{10}_{\min}(n)\right)\left((1 - q_{\mathrm{D}}(1,1)) p_0 p^{00}_{\min}(n) + (1 - q_{\mathrm{D}}(0,2)) p_1 p^{10}_{\min}(n)\right)}, \end{aligned} \tag{4.46}$$

then the discrete-time system (4.41) converges to a unique equilibrium point and the solution of (4.40) is unique.

Proof. See Appendix C.4. □

Due to the monotonicity of $p^{\theta 0}_{\min}(n)$ and $p^{\theta 0}_{\max}(n)$ shown in Lemma 4.1, c_θ decreases with n. Hence, if a given ν satisfies (4.44) for some N', then ν will satisfy (4.44) for all $n \geqslant N'$ and the equilibrium is unique. If the values of p_1, q_{D}, q_{FA}, and M are fixed, then one may deduce sufficient conditions on the value of ν to have a unique equilibrium point. See Example 4.1.

Example 4.1. *Consider $q_{FA}(2) = 0.05$, $q_D(0,2) = 0.9$, $q_D(1,1) \in \{0.5, 0.8\}$, $M \in \{4, 10\}$, and $p_1 \in [0.05, 0.5]$. One verifies whether (4.44) is satisfied considering $n = 10$ for different values of ν. One obtains that (4.44) holds if $0 < \nu \leqslant \nu_{\max}$, where $\nu_{\max}$ depends on the values of p_1, q_D, q_{FA}, and M. See Figure 4.4 for the numerical values of $\nu_{\max}$ in each case.*

4.5.3 Equilibrium point as $M \to \infty$

Both $\overline{p}^{00}$ and $\overline{p}^{10}$ can be seen as functions of M. As $M \to \infty$, Algorithm 4 turns into Algorithm 3. In this situation, if ν is properly chosen, the probabilities of false alarm and non-detection tend to zero, as shown in Proposition 4.5.

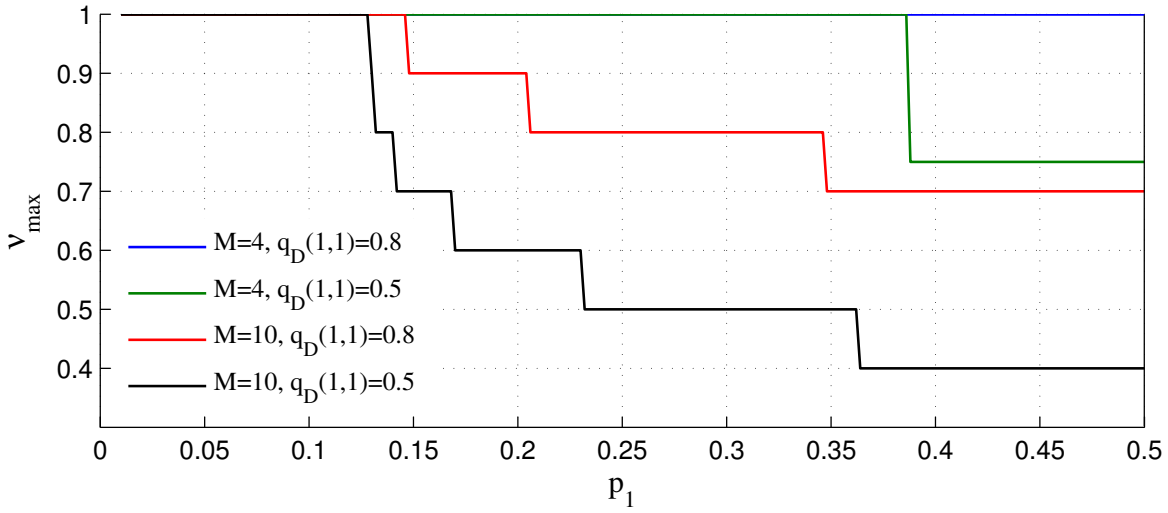

FIGURE 4.4: Upper bounds of ν to satisfy (4.44), with $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(0,2) = 0.9$, $q_{\mathrm{D}}(1,1) \in \{0.5, 0.8\}$, $M \in \{4, 10\}$, and $p_1 \in [0.05, 0.5]$.

Proposition 4.5. *If* $q_{\mathrm{FA}}(2) < \nu < q_{\mathrm{D}}(1,1)$, *then* (4.40) *has a unique solution and*

$$\lim_{M\to\infty} \overline{p}^{00} = 1, \qquad \lim_{M\to\infty} \overline{p}^{10} = 0. \tag{4.47}$$

Proof. See Appendix C.5. □

4.6 Approximations of the Equilibrium

Closed-form expressions for $\overline{p}^{00}$ and $\overline{p}^{10}$ are difficult to obtain from (4.40). This section introduces an approximation of (4.40) from which some insights may be obtained on the way ν should be chosen.

Since $\overline{p}^{10}$ represents the proportion of nodes with defective sensors that have not detected their status, the value of $\overline{p}^{10}$ should be small. From (4.35-4.36) one sees that $\lim_{\overline{p}^{10}\to 0} h_0 = q_{\mathrm{FA}}(2)$ and $\lim_{\overline{p}^{10}\to 0} h_1 = q_{\mathrm{D}}(1,1)$, thus one may consider the following approximations

$$h_0 \approx \widetilde{h}_0 = q_{\mathrm{FA}}(2), \quad h_1 \approx \widetilde{h}_1 = q_{\mathrm{D}}(1,1). \tag{4.48}$$

Therefore, (4.40) may be rewritten as

$$\begin{cases} \widetilde{p}^{00} = \sum_{k:k/M<\nu} \binom{M}{k} (q_{\mathrm{FA}}(2))^k (1 - q_{\mathrm{FA}}(2))^{M-k}, \\ \widetilde{p}^{10} = \sum_{k:k/M<\nu} \binom{M}{k} (q_{\mathrm{D}}(1,1))^k (1 - q_{\mathrm{D}}(1,1))^{M-k}. \end{cases} \tag{4.49}$$

from which one deduces approximate values $\widetilde{X}_0^{M,k}$ of $X_0^{M,k}$ at equilibrium

$$\begin{cases} \widetilde{X}_0^{M,k} = \binom{M}{k} (q_{\mathrm{FA}}(2))^k (1 - q_{\mathrm{FA}}(2))^{M-k}, \\ \widetilde{X}_1^{M,k} = \binom{M}{k} (q_{\mathrm{D}}(1,1))^k (1 - q_{\mathrm{D}}(1,1))^{M-k}. \end{cases} \tag{4.50}$$

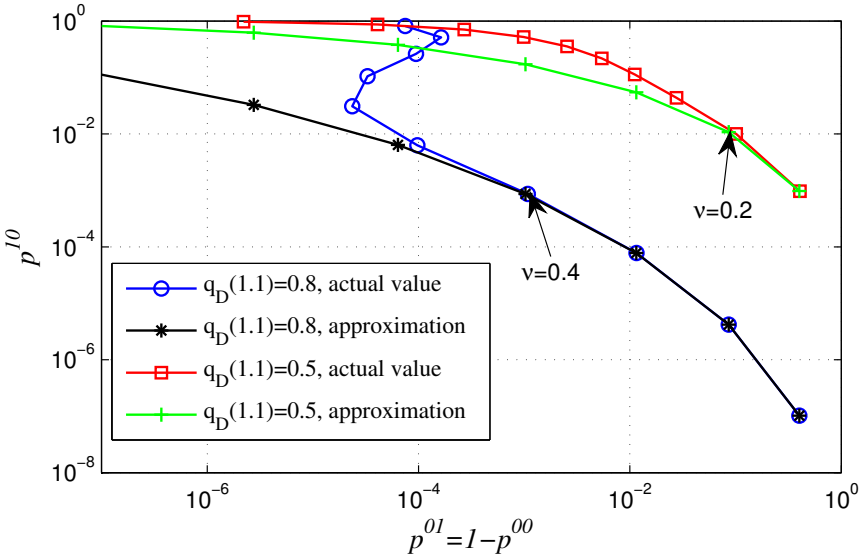

FIGURE 4.5: Approximate p^{10} as a function of approximate p^{01}, for various ν and fixed $M = 10$.

For any fixed value of M, $q_{\mathrm{FA}}(2)$, and $q_{\mathrm{D}}(1,1)$, the values of *detection rate* ($\overline{p}^{11}$) and *false alarm rate* ($\overline{p}^{01}$) at equilibrium can be predicted using (4.49), since $\overline{p}^{01} = 1-\overline{p}^{00}$ and $\overline{p}^{11} = 1-\overline{p}^{10}$. Consider for example $M = 10$, $q_{\mathrm{FA}}(2) = 0.05$, and $q_{\mathrm{D}}(1,1) = 0.8$. Figure 4.5 presents $\widetilde{p}^{10}$ as a function of $\widetilde{p}^{01}$ for different values of ν. This figure is helpful to choose the value of ν to meet different performance requirements. The actual values of $\overline{p}^{10}$ and $\overline{p}^{01}$ are also shown in Figure 4.5, which are very close to $\widetilde{p}^{10}$ and $\widetilde{p}^{01}$, in the region where $\overline{p}^{10}$ is close to 1.

4.7 Numerical results

4.7.1 Numerical verification of theoretical results

This section presents first the solution of the state equation (4.32) describing the evolution of the proportion of nodes in various states. Algorithm 4 is simulated considering a random displacement of nodes without constraint on their speed. This allows to verify the correctness of the theoretical results presented in this chapter.

Consider a LODT where $q_{\mathrm{FA}}(0,2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, and $q_{\mathrm{D}}(0,2) = 0.9$. Besides, $p_0 = 0.9$, $p_1 = 0.1$, $M = 4$, $\nu = 0.4$, and $\lambda = 1$. Figure 4.6 presents the evolution of the proportion of nodes with good sensors (top part) and defective sensors (bottom part) in different states, obtained by solving (4.32). Note that Δt represents the duration of a unit time slot used in the simulation. One observes that the proportion of nodes in each state becomes almost constant as $t/\Delta t \geqslant 15$. For the nodes with $\theta = 0$, only two states are such that $X_0^{\ell,k} > 0.05$, $(0,4,0)$ and $(0,4,1)$, while the others are very close to 0. For the nodes with $\theta = 1$, only $X_1^{4,4}$, $X_1^{4,3}$, and $X_1^{4,2}$ are relatively large. Since there is no common (ℓ,k) such that both $X_0^{\ell,k}$ and $X_1^{\ell,k}$ have unnegligeable values, the accuracy of the algorithm can be very good. With $\nu = 0.4$, one has $\overline{p}^{00} = 0.985$ and $\overline{p}^{10} = 0.027$. Only 1.5% of

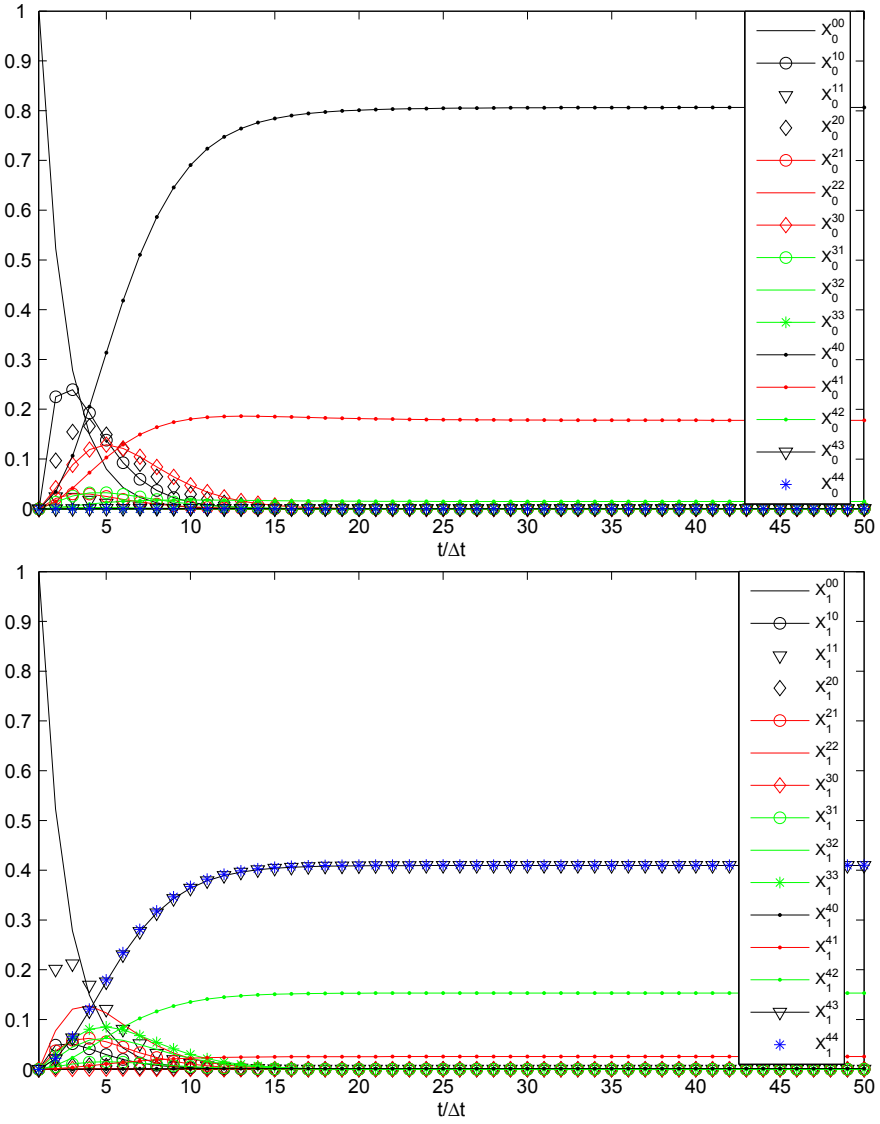

FIGURE 4.6: Evolution of $X_0^{\ell,k}(t)$ (top) and $X_1^{\ell,k}(t)$ (bottom) obtained from (4.32), in the case where $q_{\mathrm{FA}}(0,2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 4$, $\nu = 0.4$, and $\lambda = 1$.

the good nodes believe they are carrying defective sensors. Less than 3% of the nodes with defective sensors have not been detected.

Consider now a set $\mathcal{S}$ of $N_S = 1000$ moving nodes uniformly distributed over a square of unit area. In the first displacement model (jump motion model): Node i randomly chooses its location at time instant $(k+1)\Delta t$, independently from its previous location at time $k\Delta t$. Two nodes communicate only at discrete time instants $k\Delta t$ when their distance is less than r_0. Node i has its neighbors in the set $\mathcal{N}_i = \{j \in \mathcal{S} : 0 < R_{i,j} \leq r_0\}$, where $R_{i,j}$ is the distance between Nodes i and j. Furthermore, if $|\mathcal{N}_i| > 1$, we assume that Node i communicates only with its closest neighbor. Denote $\rho = \pi r_0^2 N_S$ as the average value of $|\mathcal{N}_i|$. The cardinality of $\mathcal{N}_i$ approximately follows a Poisson distribution as N_S is large enough, the inter-contact probability is thus

$$\lambda \Delta t = \mathbb{P}\{|\mathcal{N}_i| = 1\} = \rho \exp(-\rho).$$

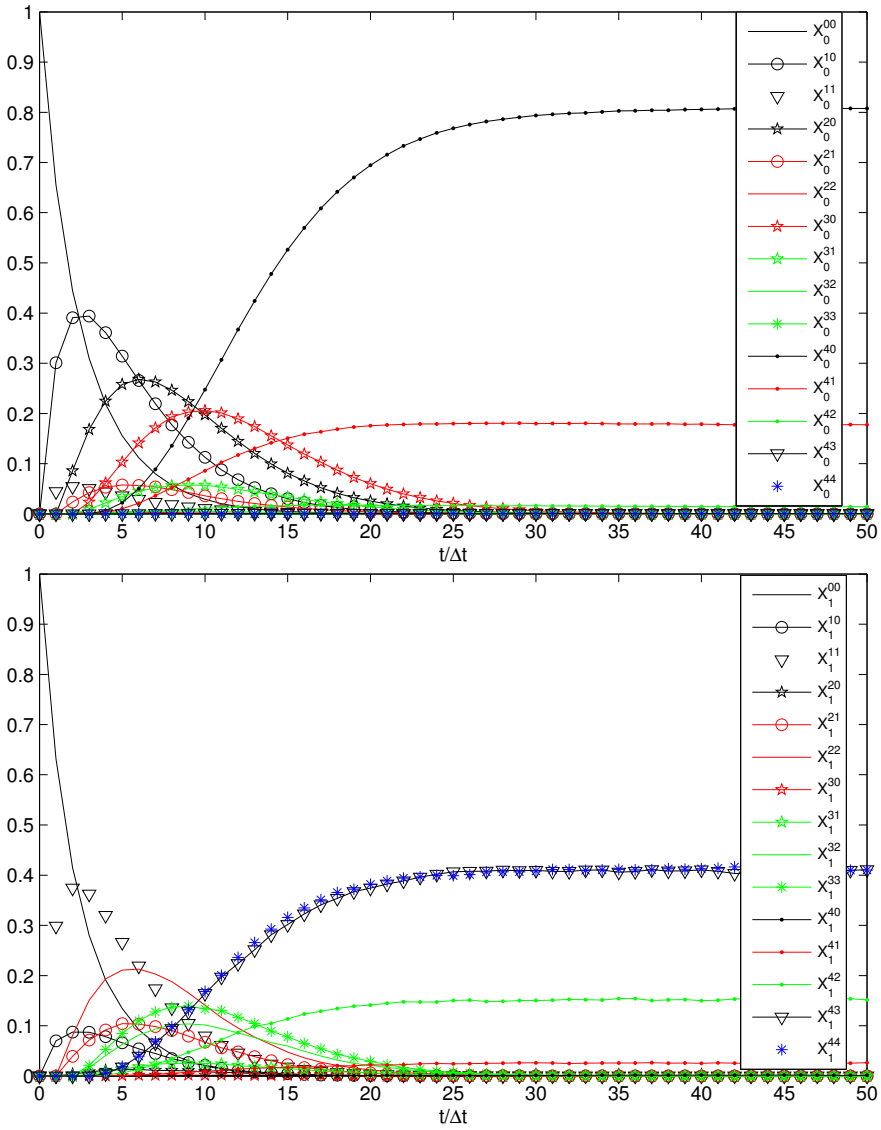

FIGURE 4.7: Evolution of $X_0^{\ell,k}(t)$ (top) and $X_1^{\ell,k}(t)$ (bottom) by simulations with the jump model, in the case where $q_{\mathrm{FA}}(0,2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 4$, $\nu = 0.4$, and $\lambda\Delta t = 0.33$.

In the Monte-Carlo simulations, we set $r_0 = 0.014$, so that $\rho \approx 0.6$ and $\lambda\Delta t \approx 0.33$. Using the same values of M, ν, q_{D}, and q_{FA} as in Figure 4.6, the simulation results for this jump motion model are shown in Figure 4.7. Comparing Figure 4.6 and Figure 4.7, one remarks that the state evolution in the transient phase has similar shape but with different convergence speed, which depends mainly on λ. Figure 4.8 shows a good match between theory and simulation for the proportions of states at equilibrium. The approximation of $\overline{X}_\theta^{4,k}$ using (4.50) is also presented in Figure 4.8, which is very close to its actual value.

4.7.2 Simulation with Brownian motion model

Consider now a Brownian motion model where each node is moving with a random speed. Each node changes its orientation when it reaches the boundary of the unit square. Define $\boldsymbol{O}_i = \left(o_x^i, o_y^i\right)$ as the

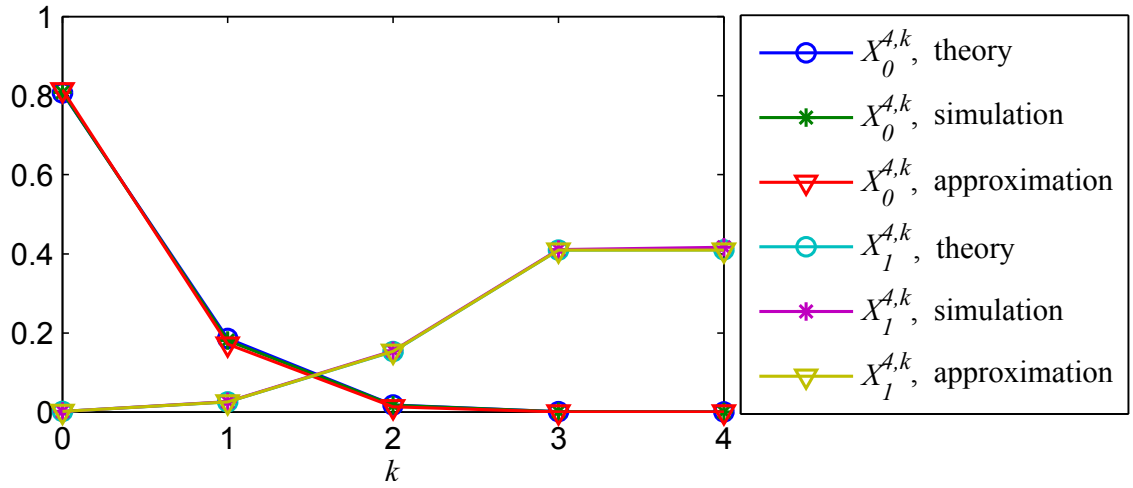

FIGURE 4.8: Comparison of $\overline{X}_{\theta}^{4,k}$ at the equilibrium.

location of Node i. Consider a second order mobility model, *i.e.*, $d^2o_x/dt^2 = v_x$ and $d^2o_y/dt^2 = v_y$, where $v_x, v_y \sim \mathcal{N}\left(0, (\sigma r_0)^2\right)$.

Consider $\sigma \in \{0.1, 1\}$, $q_{\text{FA}}(2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$, $M = 10$, and $\nu = 0.4$. Figure 4.9 compares the evolution of p^{01} and p^{10} as functions of time for the jump motion model and the Brownian motion model, with fixed $\rho \approx 0.6$. At equilibrium, the performance obtained for both models is quite close. However, the convergence speed depends on the inter-contact rate λ. When $\sigma = 0.1$, the algorithm converges slowly in the Brownian motion model. When $\sigma = 1$, which results to a larger value of λ, the evolution of p^{01} and p^{10} with the Brownian motion model are close to the jump motion model.

At the beginning of the algorithm, each node believes that its sensors are good, thus $p^{01}(0) = 0$ and $p^{10}(0) = 1$. During the algorithm, $p^{10}(t)$ decreases in the transient phase until it reaches the equilibrium. Whereas, $p^{01}(t)$ increases at first and then decreases to the equilibrium. This comes from the fact that $p^{10}(t)$ is large at the beginning and the LODT performed on a good node often detects outliers.

4.7.3 Simulation with real databases

In this section, Algorithm 4 is executed on some experimental databases instead of motion models. These databases, provided by the MIT Reality Mining Project [EP06] and the Haggle Project [Sco+09], are well investigated in several previous works, *e.g.*, [HCY11]. In this work, we use the following databases:

- *Reality*, where $N_S = 97$, lasts more than 200 days with about 111 inter-contacts per day.
- *Infocom05*, where $N_S = 41$, lasts 3 days with approximately 312 inter-contacts every hour.

More specifically, one is interested in the inter-contact trace, *i.e.*, which pair of nodes have a meeting at which time. The traces were taken from [Orl], which are converted from the original databases [EP06; Sco+09].

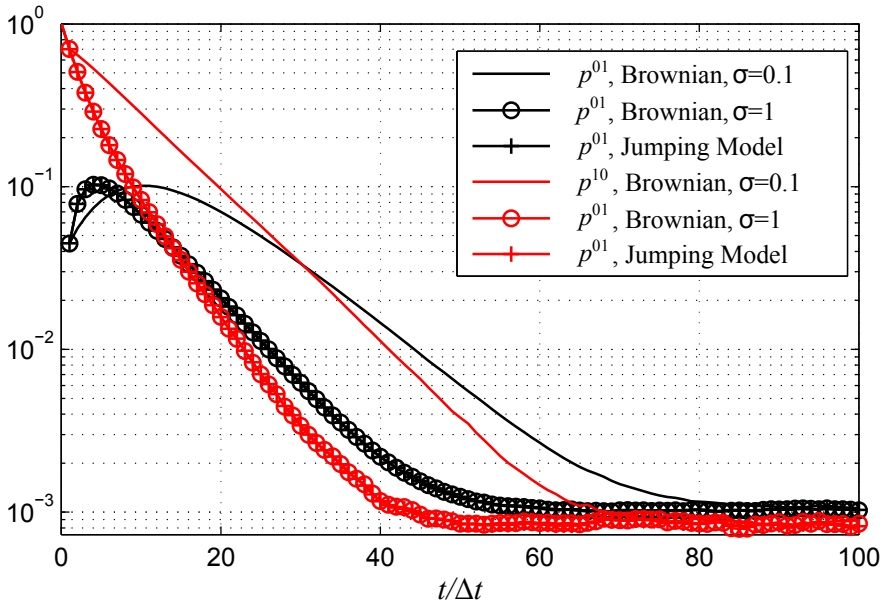

FIGURE 4.9: Evolution of p^{01} and p^{10} for the two moving models, with $\sigma \in \{0.1, 1\}$, $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$ and $\nu = 0.4$.

Consider again the following parameters: $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$, and $\nu = 0.4$. Monte-Carlo simulations are performed 500 times for each database. In each test, N_D nodes with random index are chosen to be defective. One sets $N_D = 10$ in *Infocom05* and $N_D = 20$ in *Reality*. At the top of Figures 4.10 - 4.11, the index of the active nodes (which have contact with the others) are presented at each time to show the frequency of the inter-contacts at different epochs. The evolution of p^{10} and p^{01} is plotted at the bottom of Figures 4.10 - 4.11. Interestingly, both p^{10} and p^{01}, obtained by both databases, decrease to 10^{-3} after a sufficient long time. One also observes that the convergence speed of p^{10} and p^{01} is highly related to the inter-contact rate (reflected by the density of points in the sub-figures at the top): variations are significant at beginning of working hours.

Figure 4.12 represents the states at equilibrium $\overline{X}_{\theta}^{M,k}$ obtained by using the databases *Reality* and *Infocom05*, and also by the approximation (4.50). There is an excellent match between the values at equilibrium predicted by theory and those obtained in practice.

4.7.4 Influence of the parameters

This section characterizes the influence of the parameters, such as p_1, $q_D(1,1)$, and M, on the performance of Algorithm 4. The jump motion model is used throughout this section to describe the displacement of the nodes.

Consider fixed $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, the evolution of p^{10} and p^{01} for various $p_1 \in \{0.1, 0.5\}$ and $M \in \{4, 10, 20\}$ is shown in Figure 4.13. For each different case, the value of ν is chosen such that it minimizes $\widetilde{p}^{01} + \widetilde{p}^{10}$. One observes that a large M leads to a better performance at equilibrium. The price to be paid is a longer time required to reach equilibrium. When $M = 10$, both $\overline{p}^{10}$ and $\overline{p}^{01}$ are around 10^{-3}. The proportion of the nodes with defective

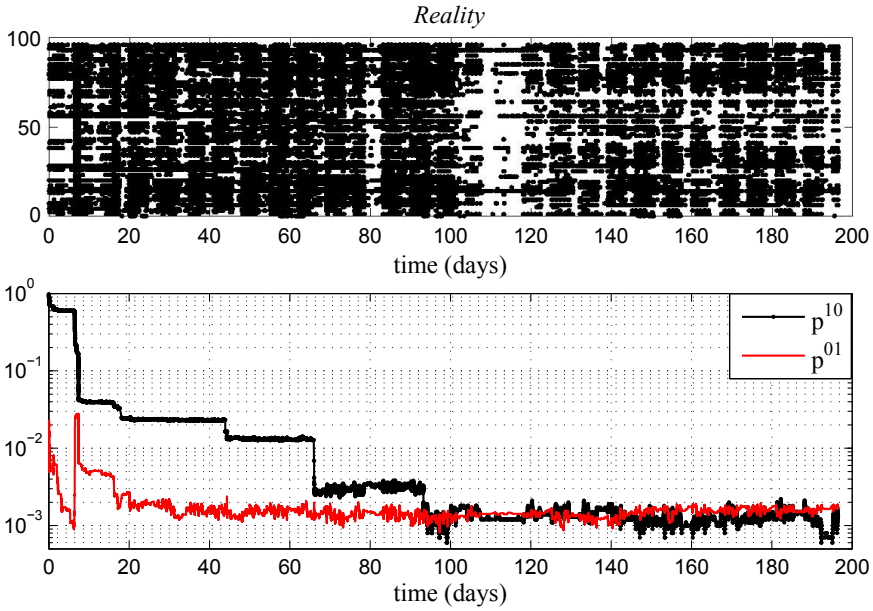

FIGURE 4.10: Indexes of active nodes (having met another node) at different time (top) and evolution of p^{01} and p^{10} (bottom) obtained by using the *Reality* database, with $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$, and $\nu = 0.4$.

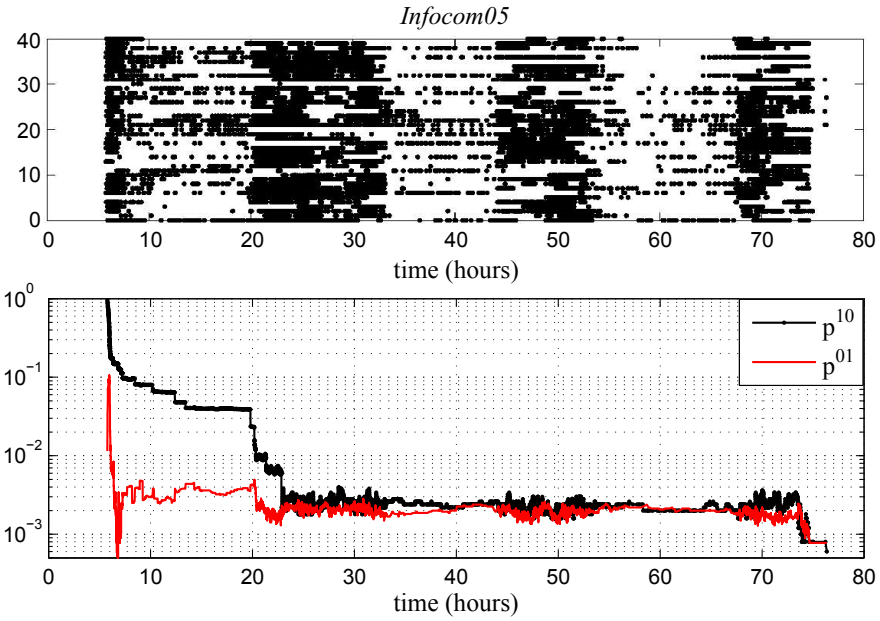

FIGURE 4.11: Indexes of active nodes at different time (top) and evolution of p^{01} and p^{10} (bottom) obtained by using the *Infocom05* database, with $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$, and $\nu = 0.4$.

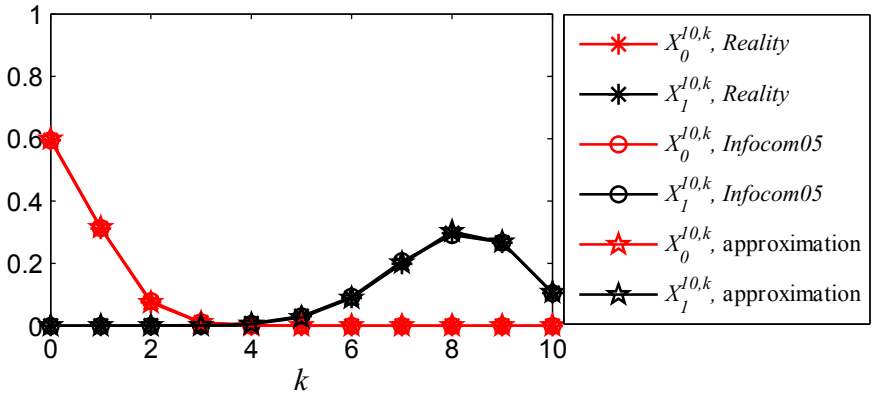

FIGURE 4.12: Comparison of $\overline{X}_{\theta}^{10,k}$ at the equilibrium obtained using the *Reality* database, the *Infocom05* database, and predicted by the approximation (4.50).

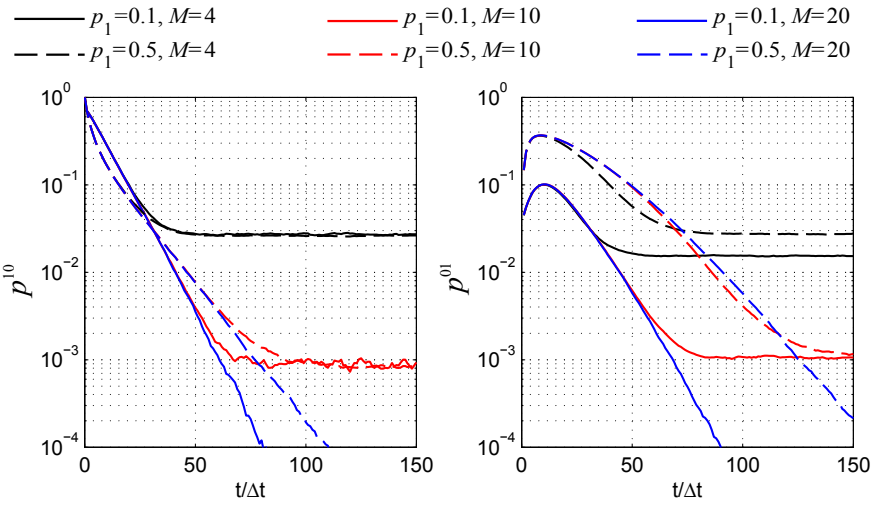

FIGURE 4.13: Evolution of p^{10} and p^{01} for various $M \in \{4, 10, 20\}$ and $p_1 \in \{0.1, 0.5\}$, with $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$.

sensors has also an impact on the convergence speed of the algorithm. For example, when p_1 is large, more time is needed to achieve a given level of performance (in terms of $\overline{p}^{10}$ and $\overline{p}^{01}$).

To show the effectiveness of the proposed DTN-DFD algorithm, consider now $q_{\mathrm{D}}(0,2) = 0.9$ and $M = 10$. For $p_1 = 0.1$ and $p_1 = 0.5$, one is interested in the achievable $\overline{p}^{10}$ and $\overline{p}^{01}$ for $0 \leqslant q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(0,2)$ and $q_{\mathrm{FA}}(2) < q_{\mathrm{D}}(1,1) \leqslant q_{\mathrm{D}}(0,2)$. Four areas are considered:

- Area 3: both $\overline{p}^{10}$ and $\overline{p}^{01}$ are less than 10^{-3} ;
- Area 2: both $\overline{p}^{10}$ and $\overline{p}^{01}$ are less than 10^{-2} ;
- Area 1: both $\overline{p}^{10}$ and $\overline{p}^{01}$ are less than 10^{-1};
- Area 0: either $\overline{p}^{10}$ or $\overline{p}^{01}$ cannot be less than 10^{-1}.

Figure 4.14 shows partition of the $(q_{\mathrm{D}}(1,1), q_{\mathrm{FA}}(2))$ triangle in four areas, represented in different colors. The ratio of defective nodes in the network has not a significant impact on the performance at the equilibrium, even when 50% of nodes are defective.

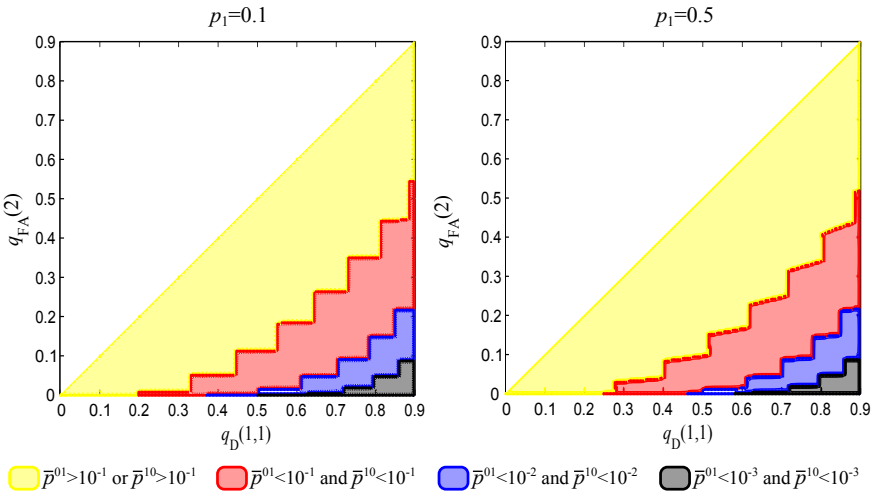

FIGURE 4.14: Achievable $\overline{p}^{10}$ or $\overline{p}^{01}$ for different values of the pair $(q_D(1,1), q_{FA}(2))$ when $p_1 = 0.1$ (left) and for $p_1 = 0.5$ (right).

4.7.5 Experiment

The following experiment has been conducted at the EuWIn platform at University of Bologna. $N_S = 35$ wireless sensor nodes with unique identifier i have been used. The transmission range of each node is tuned approximately at 1 m. Each node maintains a local timer and broadcasts, at a period $\Delta t = 1$s , a packet containing its identity. Each time a node receives a packet, it records the identity of the encountered node as well as the current value of the local timer.

During the break of a course (which lasts about 15 minutes), the nodes are distributed 34 students. At the beginning of the experiment, the nodes were close to each other so that a master node can be used to make all the nodes start their timer at the same time. The students were asked to move freely in the university and then to return the nodes at the end of the break. The proposed algorithm has then been executed on the obtained traces with the same parameters as in Section 4.7.3. In each of the 1000 simulations, $N_D = 5$ nodes with random index are chosen to be defective. Similar to Figure 4.10 - 4.11, the results of post-processing are shown in Figure 4.15. Both p^{10} and p^{01} are decrease to less than 10^{-2} . One also observes that variations are significant at the end, as the students were back to class.

4.8 Conclusion

This chapter presents a fully distributed algorithm allowing each node of a DTN to estimate the status of its own sensors using LODT performed during the meeting of nodes. The DTN-DFD algorithm is analyzed considering a Markov model of the evolution of the proportion of nodes with a given belief in their status. This model is then used to derive the evolution of the proportions of the nodes in different states. The existence and uniqueness of an equilibrium is discussed. Interestingly, the proportions at the equilibrium follow a Binomial distribution. Approximations of these proportions of nodes provide a clear guide to properly choose the decision parameter of the DTN-DFD algorithm. In the simulations, a jump motion model and a Brownian motion model are

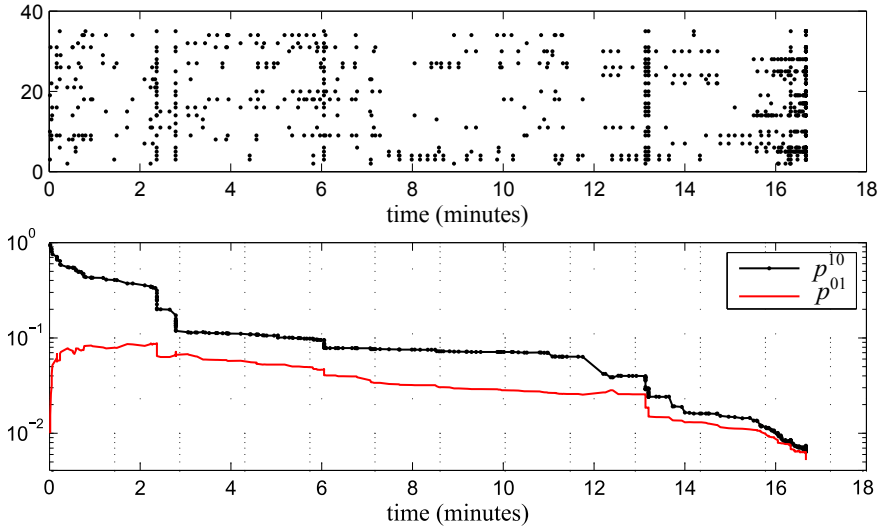

FIGURE 4.15: Indexes of active nodes at different time (top) and evolution of p^{01} and p^{10} (bottom) obtained by using *our trace*, with $q_{\mathrm{FA}}(2) = 0.05$, $q_{\mathrm{D}}(1,1) = 0.8$, $q_{\mathrm{D}}(0,2) = 0.9$, $M = 10$, and $\nu = 0.4$.

considered. The results show a good match with theory. The convergence speed of the DTN-DFD algorithm depends on the inter-contact rate and on the proportion of nodes with defective sensors p_1. Nevertheless, p_1 has not a significant impact on the non-detection and false alarm rates at equilibrium, showing the robustness of the approach to a large number of defective nodes.

Further research will be dedicated to an adaptation of ν with time to increase the convergence speed of the proposed DTN-DFD algorithm. This may be particularly important in variants of the considered problem, such as malware detection.

Chapter 5

Distributed Self-Rating by Pairwise Interaction

Consider a community of agents [1] that perform some task (estimation, detection...) with different levels of ability (LoA). As indicated in Section 1.1.2, knowing the LoA of each agent has a lot of benefits, *e.g.*, facilitating the measurement processing in the context of wireless sensor networks. This chapter presents and analyzes a distributed algorithm to make each agent in the community assess its own LoA. Several related work has been presented in Section 1.2.4.

The chapter is organized as follows. Section 5.1 introduces the system model and the local comparison test (LCT). Section 5.2 describes and analyzes the proposed DSR algorithm. Section 5.3 discusses the property of the equilibrium obtained from the dynamic state equations describing the evolution with time of the proportion of agents with a given belief in their expertise. Simulations results are reported in Section 5.4 before drawing some conclusions in Section 5.5.

5.1 System Model and Local Comparison Test

Consider a set $\mathcal{A}$ of N_{A} moving agents. The status $\theta_i \in \Theta = \{1, \ldots, K\}$ of Agent i, assumed constant with time, describes its LoA at performing some task. Based on the different status of the agents, $\mathcal{A}$ can be partitioned into K groups denoted $\mathcal{A}_1, \ldots, \mathcal{A}_K$, with $\mathcal{A}_\theta = \{i \in \mathcal{A} : \theta_i = \theta\}$. Denote p_θ as the proportion of the agents belonging to $\mathcal{A}_\theta$.

Without loss of generality, we assume that groups are sorted in decreasing LoA of their agents. Thus, agents in $\mathcal{A}_1$ are the best-performing and those in $\mathcal{A}_K$ are the worst-performing. Initially, Agent i is not aware of the actual value of θ_i but is willing to estimate it as fast as possible. As in Chapter 4, one considers the following assumptions on the agents

- *A1*) $\theta_i(t) = \theta_i$, *i.e.*, the group to which Agent i belongs does not change over the time horizon of the experiment;
- *A2*) only pairwise interactions between agents are considered;
- *A3*) the agents are mobile and form a well-mixed population;
- *A4*) the time interval between two successive meetings follows an exponential distribution with an inter-contact rate λ [Zhu+10].

[1] Notice that "agent" is a more general conception than "node".

In what follows, one will present a DSR algorithm allowing each Agent i to get an estimate $\widehat{\theta}_i$ of its status θ_i. For that purpose an LCT is introduced, performed by Agent i after meeting and interacting with an other Agent j. The output $y_{i,j}$ of the LCT performed by Agent i upon interaction with Agent j is binary. The LCT yields $y_{i,j} = 1$ when it is likely that Agent i is not worse at doing the considered task than Agent j. It yields $y_{i,j} = 0$ when Agent i is worse than Agent j. The LCT is unable to determine the absolute LoA of Agent i. LCTs may provide erroneous conclusions and are characterized by their statistical properties

$$q(\theta_i, \theta_j) = \mathbb{P}\{Y_{i,j} = 1 | \theta_i, \theta_j\}. \tag{5.1}$$

As will be seen in Section 5.4, one has not necessarily $q(\theta_i, \theta_j) = 1 - q(\theta_j, \theta_i)$. Concerning the LCT, one considers the following assumptions

- *A5)* $\theta_i \leqslant \theta_j$ implies $q(\theta_i, 1) \geqslant q(\theta_j, 1)$ and $q(1, \theta_i) \geqslant q(1, \theta_j)$,
- *A6)* for all $(\theta_i, \theta_j) \in \Theta^2$,one has $q(\theta_i, \theta_j) > 0$.

A5) appears reasonable, since the first group contains the agents with the highest LoA, and, compared to an agent of intermediate expertise, an agent with a low LoA has less probability to be deemed better than the best-performing agents. A6) means that even if θ_i corresponds to the group with the worst agents, the LCT has a non-zero probability to conclude that Agent i is better than Agent j. This probability can be made arbitrarily small, but is required to show the existence of an equilibrium of the dynamics describing the evolution of the proportions of Agents with a given belief in their status, see Section .

Interactions may take various forms depending on the application scenario, ranging from the exchange of noisy measurements m_i and m_j of the same physical quantity when the agents are nodes of a WSN, to a blitz-game between humans, willing, *e.g.*, to self-rate their LoA. An example LCT is provided in Section 5.4 in the context of WSNs. A meeting between two nodes does not necessarily entail interaction: one assumes here that the probability of interaction $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ is a function of the estimates $\widehat{\theta}_i(t)$ and $\widehat{\theta}_j(t)$ of the groups to which each agent belongs. When the agents aim is to self-rate their LoA in doing some task, for instance, $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ will approach one when $\widehat{\theta}_i$ and $\widehat{\theta}_j$ are close and thus the outcome of the LCT is difficult to predict. When $\widehat{\theta}_i$ and $\widehat{\theta}_j$ are very different, $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ may be small, thus preventing the agents wasting time in a LCT whose outcome is reputed easily foreseeable.

In practice, one assumes that when two agents i and j meet, they start exchanging their estimated status, Agent i will request for a further interaction with probability $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ and Agent j with a probability $\alpha\left(\widehat{\theta}_j, \widehat{\theta}_i\right)$. Interaction occurs when at least one of the agents request it. Only agents requesting an interaction perform the LCT and exploit its result. The values of $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$ are design parameters that can be adjusted to optimize the performance of the DSR algorithm.

5.2 Distributed Self-Rating Algorithm

In the proposed DSR algorithm, each agent manages two counters $c_{\mathrm{t},i}(t)$ and $c_{\mathrm{b},i}(t)$ initialized at 0 at $t = 0$. The number of LCT performed by Agent i following an interaction it has requested is stored in $c_{\mathrm{t},i}(t)$. The number of tests concluding that Agent i is *better* than the agent met is stored in $c_{\mathrm{b},i}(t)$. As indicated in Section 5.1, an agent involved in an interaction it has not requested does not update $c_{\mathrm{t},i}(t)$ and $c_{\mathrm{b},i}(t)$. If the agents are randomly spread and moving (Assumption $A3$), the ratio $c_{\mathrm{b},i}(t)/c_{\mathrm{t},i}(t)$ will mainly depend on the proportions of agents in each group, on the interaction probabilities, and on the statistical properties of the LCT.

Intuitively, an agent with status θ is likely to have a larger ratio $c_{\mathrm{b},i}(t)/c_{\mathrm{t},i}(t)$ than an agent with status θ' when $\theta < \theta'$. One may thus introduce a partition of the interval $[0, 1]$ into K decision intervals $[\nu_k, \nu_{k-1})$ with $\nu_0 = 1$ and $\nu_K = 0$ and consider the decision rule

$$\widehat{\theta}_i\,(t) = k \text{ if } c_{\mathrm{b},i}(t)/c_{\mathrm{t},i}(t) \in [\nu_k, \nu_{k-1})\,,\ k = 1\ldots, K. \tag{5.2}$$

The aim of this work is to determine conditions on $p_1, \ldots, p_K$ and q to show that the decision rule (5.2) leads to a satisfying self-rating of the agents, for appropriate choice of the value of ν_k, for all $k = 1, \ldots, K$.

5.2.1 Practical Self-Rating Algorithm

Let $\mathbf{x}_i(t) = (\theta_i, c_{\mathrm{t},i}(t), c_{\mathrm{b},i}(t))$ represent the *state* of each Agent i. If all the LCT results obtained in the past are considered, then $c_{\mathrm{t},i}(t)$ can go unbounded, which results in an infinite number of possible values for $\mathbf{x}_i(t)$. The global behavior of the algorithm is in this case difficult to analyze. To limit the number of possible states, one considers the evolution of $c_{\mathrm{t},i}(t)$ and $c_{\mathrm{b},i}(t)$ over a sliding variable-length time interval containing the time instants of the last M meetings during which Agent i has performed a LCT. Algorithm 5 summarizes the proposed DSR algorithm for an arbitrary reference Agent i.

5.2.2 Macroscopic evolution

At time t, among the agents with status θ, let $X_\theta^{\tau,\beta}\,(t)$ be the proportion of agents in state $\mathbf{x}_i(t) = (\theta, \tau, \beta)$, *i.e.*, with $c_{\mathrm{t},i}\,(t) = \tau$, $c_{\mathrm{b},i}\,(t) = \beta$, and $0 \leqslant \beta \leqslant \tau \leqslant M$. The evolution of the state of Agent i with status θ_i, follows a Markov model with state transition diagram similar to that considered in DTN-DFD. With the initial conditions $X_\theta^{0,0}\,(0) = 1$ and $X_\theta^{\tau,\beta}\,(0) = 0$, $\forall \tau, \beta \neq 0$, the evolution of the various proportions $X_\theta^{\tau,\beta}(t)$ of agents in the corresponding states, with $\theta = 1, \ldots, K$

Algorithm 5 DSR algorithm for Agent i

1. Initialize $t_i^0 = 0$, $\widehat{\theta}_i(0) = 1$, $c_{\mathrm{t},i}(0) = c_{\mathrm{b},i}(0) = 0$, $\kappa = 1$, and $\mu = 0$.
2. Do $\widehat{\theta}_i(t) = \widehat{\theta}_i\left(t_i^{\kappa-1}\right)$, $c_{\mathrm{t},i}(t) = c_{\mathrm{t},i}\left(t_i^{\kappa-1}\right)$, $c_{\mathrm{b},i}(t) = c_{\mathrm{b},i}\left(t_i^{\kappa-1}\right)$, and $t = t + \delta t$ until the κ-th meeting occurs at time t_i^κ with Agent $j \in \mathcal{A}$.
3. Transmit $\widehat{\theta}_i\left(t_i^\kappa\right)$ to Agent j^κ and receive $\widehat{\theta}_j\left(t_i^\kappa\right)$ from Agent j^κ.
4. With probability $\alpha\left(\widehat{\theta}_i\left(t_i^\kappa\right), \widehat{\theta}_j\left(t_i^\kappa\right)\right)$, perform a LCT with outcome y_i^μ, then
   (a) $\mu = \mu + 1$. Update $c_{\mathrm{t},i}$ and $c_{\mathrm{b},i}$ as
   $$\begin{cases} c_{\mathrm{t},i}(t_i^\kappa) = \min\{\mu, M\} \\ c_{\mathrm{b},i}(t_i^\kappa) = \sum_{m=\max\{1,\mu-M+1\}}^{\mu} y_i^m \end{cases} \tag{5.3}$$
   (b) Update $\widehat{\theta}_i$ according to (5.2)
5. $\kappa = \kappa + 1$.
6. Go to 2.

and $0 < \beta \leqslant \tau < M$, are described by

$$\begin{cases} \frac{dX_\theta^{0,0}}{dt} \overset{(a)}{=} -\lambda X_\theta^{0,0}\left(\pi_\theta^{1,0}(0,0) + \pi_\theta^{1,1}(0,0)\right), \\ \frac{dX_\theta^{\tau,0}}{dt} \overset{(b)}{=} \lambda\left(-X_\theta^{\tau,0}\left(\pi_\theta^{1,0}(\tau,0) + \pi_\theta^{1,1}(\tau,0)\right) + X_\theta^{\tau-1,0}\pi_\theta^{1,0}(\tau-1,0)\right), \\ \frac{dX_\theta^{\tau,\tau}}{dt} \overset{(c)}{=} \lambda\left(-X_\theta^{\tau,\tau}\left(\pi_\theta^{1,0}(\tau,\tau) + \pi_\theta^{1,1}(\tau,\tau)\right) + X_\theta^{\tau-1,\tau-1}\pi_\theta^{1,1}(\tau-1,\tau-1)\right), \\ \frac{dX_\theta^{M,0}}{dt} \overset{(d)}{=} \lambda\left(-X_\theta^{M,0}\pi_\theta^{0,1}(M,0) + X_\theta^{M-1,0}\pi_\theta^{1,0}(M-1,0) + X_\theta^{M,1}\pi_\theta^{0,-1}(M,1)\right), \\ \frac{dX_\theta^{M,M}}{dt} \overset{(e)}{=} \lambda\left(-X_\theta^{M,M}\pi_\theta^{0,-1}(M,M) + X_\theta^{M,M-1}\pi_\theta^{0,1}(M,M-1) + X_\theta^{M-1,M-1}\pi_\theta^{1,1}(M-1,M-1)\right), \\ \frac{dX_\theta^{\tau,\beta}}{dt} \overset{(f)}{=} \lambda\left(-X_\theta^{\tau,\beta}\left(\pi_\theta^{1,0}(\tau,\beta) + \pi_\theta^{1,1}(\tau,\beta)\right) + X_\theta^{\tau-1,\beta}\pi_\theta^{1,0}(\tau-1,\beta) + X_\theta^{\tau-1,\beta-1}\pi_\theta^{1,1}(\tau-1,\beta-1)\right) \\ \frac{dX_\theta^{M,\beta}}{dt} \overset{(g)}{=} \lambda\Big(-X_\theta^{M,\beta}\left(\pi_\theta^{0,1}(M,\beta) + \pi_\theta^{0,-1}(M,\beta)\right) + X_\theta^{M-1,\beta-1}\pi_\theta^{1,1}(M-1,\beta-1) \\ \qquad\qquad + X_\theta^{M-1,\beta}\pi_\theta^{1,0}(M-1,\beta) + X_\theta^{M,\beta-1}\pi_\theta^{0,1}(M,\beta-1) + X_\theta^{M,\beta+1}\pi_\theta^{0,-1}(M,\beta+1)\Big), \end{cases} \tag{5.4}$$

in which $\pi_\theta^{\delta_\mathrm{t},\delta_\mathrm{b}}$ denotes the transition probability from State (θ,τ,β) to State $(\theta,\tau+\delta_\mathrm{t},\beta+\delta_\mathrm{b})$. Although the state equations are the same as those in DTN-DFD, $\pi_\theta^{\delta_\mathrm{t},\delta_\mathrm{b}}$ differs due to the increased number of groups and to the interaction probability α.

Note that $\pi_\theta^{\delta_\mathrm{t},\delta_\mathrm{b}}$ depends on the current state (θ,τ,β) of the reference Agent i, but also on the current proportion of agents with estimated status $\widehat{\theta}_i(t)$. One has first to evaluate the probability that some Agent i with state (θ,τ,β), *i.e.*, estimated status $\widehat{\theta}(\tau,\beta)$ performs a LCT during a meeting with a random Agent J. This probability may be evaluated as

$$\gamma(\tau,\beta) = E\left(\alpha\left(\widehat{\theta}(\tau,\beta), \widehat{\theta}_J(t)\right)\right) \tag{5.5}$$

where the expectation has to be taken over $\widehat{\theta}_J(t)$. Then

$$\begin{aligned}\gamma(\tau,\beta) &= \sum_{k_1,k_2\in\Theta} \alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)\mathbb{P}\left\{\widehat{\theta}_J(t)=k_2,\theta_J=k_1\right\}\\ &= \sum_{k_1,k_2\in\Theta} \alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)\mathbb{P}\left\{\widehat{\theta}_J(t)=k_2|\theta_J=k_1\right\}\mathbb{P}\{\theta_J=k_1\}\\ &= \sum_{k_1,k_2\in\Theta} \alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)p_{k_1}p^{k_1k_2}(t),\end{aligned} \tag{5.6}$$

where $p^{k_1k_2}(t)=\mathbb{P}\left\{\widehat{\theta}_J(t)=k_2|\theta_J=k_1\right\}$ is the proportion of agents with status k_1 believing their status is k_2. Using (5.2), one deduces that

$$\begin{aligned}p^{k_1k_2}(t) &= \mathbb{P}\left\{\widehat{\theta}_J(t)=k_2|\theta_J=k_1\right\}\\ &= \begin{cases} X_{k_1}^{0,0}(t)+\sum_{\tau,\beta,\beta/\tau\in[\nu_{k_2},\nu_{k_2-1})}X_{k_1}^{\tau,\beta}(t), & \text{if } k_2=1,\\ \sum_{\tau,\beta,\beta/\tau\in[\nu_{k_2},\nu_{k_2-1})}X_{k_1}^{\tau,\beta}(t), & \text{else.}\end{cases}\end{aligned} \tag{5.7}$$

For each agent, two phases have to be considered in Algorithm 5, depending on the value of $c_{\mathrm{t},i}(t)$. In the *transient regime,* for states with $c_{\mathrm{t},i}(t)=\tau<M$, one has $(\delta_\mathrm{t},\delta_\mathrm{b})\in\{(0,0),(1,0),(1,1)\}$, since τ may either increase or remain constant and $\delta_\mathrm{b}\leqslant\delta_\mathrm{t}$. The only possibility leading to $\delta_\mathrm{t}=0$ is that Agent i, once it has met a random Agent J, decides to interact. Then

$$\pi_\theta^{0,0}(t,\tau,\beta) \;=\; 1-\gamma(\tau,\beta). \tag{5.8}$$

A state transition occurs with $(\delta_\mathrm{t},\delta_\mathrm{b})=(1,1)$ when, once Agent i has met Agent J, they continue interacting *and* the LCT yields $y_{i,J}(t)=1$. Since α only depends on the estimates of the status, these two events can be assumed as independent and one has to consider all possible values taken by $\widehat{\theta}_J(t)$ to get

$$\begin{aligned}&\pi_\theta^{1,1}(t,\tau,\beta)\\ &= \sum_{k_2\in\Theta}\alpha\left(\widehat{\theta}(\tau,\beta),\widehat{\theta}_J(t)=k_2\right)\mathbb{P}\left\{Y_i=1,\widehat{\theta}_J(t)=k_2|\theta_i=\theta\right\}\\ &= \sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)\mathbb{P}\left\{Y_i=1,\theta_J=k_1,\widehat{\theta}_J(t)=k_2|\theta_i=\theta\right\}\\ &= \sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)\mathbb{P}\{Y_i=1|\theta_i=\theta,\theta_J=k_1\}\,\mathbb{P}\left\{\widehat{\theta}_J(t)=k_2|\theta_J=k_1\right\}\mathbb{P}\{\theta_J=k_1\}\\ &= \sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)p_{k_1}p^{k_1k_2}(t)\,q(\theta,k_1).\end{aligned} \tag{5.9}$$

Then, $\pi_\theta^{1,0}(t,\tau,\beta)$ is obtained similarly

$$\pi_\theta^{1,0}(t,\tau,\beta)=\sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(\tau,\beta),k_2\right)p_{k_1}p^{k_1k_2}(t)\left(1-q(\theta,k_1)\right). \tag{5.10}$$

In the *permanent regime*, one can obtain the following transition probabilities using the similar derivations as considered in Section 4.3.2,

$$\pi_\theta^{0,1}(t,M,\beta)=\sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(M,\beta),k_2\right)p_{k_1}p^{k_1k_2}(t)\,q(\theta,k_1)\,\frac{M-\beta}{M}, \tag{5.11}$$

$$\pi_\theta^{0,-1}(t,M,\beta)=\sum_{k_1,k_2\in\Theta}\alpha\left(\widehat{\theta}(M,\beta),k_2\right)p_{k_1}p^{k_1k_2}(t)\,(1-q(\theta,k_1))\,\frac{\beta}{M}. \tag{5.12}$$

and $\pi_\theta^{0,0}(t,M,\beta)=1-\pi_\theta^{0,1}(t,M,\beta)-\pi_\theta^{0,-1}(t,M,\beta)$.

5.3 Analysis of Equilibrium

In this section, the asymptotic behavior of the state equations (5.4) is characterized. Let $\overline{X}_\theta^{\tau,\beta}$ be the value at equilibrium of $X_\theta^{\tau,\beta}$. The proportion of agents with status θ estimating their status as $\widehat{\theta}$ depends on the partition of the interval $[0,1]$ introduced in (5.2)

$$\overline{p}^{\theta\widehat{\theta}}=\sum_{\tau=1}^{M}\sum_{\beta:\beta/\tau\in\left[\nu_{\widehat{\theta}},\nu_{\widehat{\theta}-1}\right)}\overline{X}_\theta^{\tau,\beta}. \tag{5.13}$$

This analysis for general αs is challenging, one thus considers the following two special cases.

- *Case I*, where a meeting always leads to an interaction, *i.e.*,

$$\alpha(k_1,k_2)=1,\qquad\forall k_1,k_2. \tag{5.14}$$

- *Case II*, where an interaction is only performed when an Agent i meets an Agent j believing its status is 1, *i.e.*,

$$\alpha(k_1,k_2)=\begin{cases}1 & \text{if } k_2=1\\ 0 & \text{else.}\end{cases} \tag{5.15}$$

5.3.1 Case I

To lighten the notations, introduce

$$s_\theta=\sum_{k\in\Theta}p_k q(\theta,k), \tag{5.16}$$

then $\pi_\theta^{\delta_t,\delta_b}$ introduced in Section 5.2.2 may be rewritten as

$$\begin{cases}\pi_\theta^{1,1}(t,\tau,\beta)=s_\theta, & \pi_\theta^{0,1}(t,M,\beta)=\frac{M-\beta}{M}s_\theta,\\ \pi_\theta^{1,0}(t,\tau,\beta)=1-s_\theta, & \pi_\theta^{0,-1}(t,M,\beta)=\frac{\beta}{M}(1-s_\theta).\end{cases} \tag{5.17}$$

Clearly, the transition probabilities is now time-invariant.

Proposition 5.1. *In Case I, the dynamic system (5.4) admits an equilibrium* $\overline{X}_\theta^{\tau,\beta}$*, for all* $\theta \in \Theta$ *and* $d \leqslant \tau$*, with*

$$\overline{X}_\theta^{\tau,\beta} = \begin{cases} 0, & \forall \tau < M, \\ \binom{M}{\beta} (s_\theta)^\beta (1-s_\theta)^{M-\beta}, & \tau = M. \end{cases} \tag{5.18}$$

Proof. Consider a reference Agent i with $c_{\mathrm{t},i} = \tau < M$. In Case I, from (5.17), one has $\pi_\theta^{1,0}(t,\tau,\beta) + \pi_\theta^{1,1}(t,\tau,\beta) = 1$. As a consequence, when Agent i meets an other agent, $c_{\mathrm{t},i}$ increases. Hence, $c_{\mathrm{t},i}$ will reach M after M meetings and as $t \to \infty$, no node will be in a state (θ,τ,β) with $\tau < M$. Hence, necessarily $\overline{X}_\theta^{\tau,\beta} = 0$, for all $\tau < M$ and $\beta \leqslant \tau$. The derivation is similar to Appendix C.2 to prove the distribution of $\overline{X}_\theta^{M,\beta}$. □

With the closed-form expression of $\overline{X}_\theta^{\tau,\beta}$ presented in Proposition 5.1, one gets the *correct decision rate* (CDR) of agents with status $\theta \in \Theta$,

$$\overline{p}^{\theta\theta} = \sum_{\beta:\beta/M \in [\nu_\theta, \nu_{\theta-1})} \binom{M}{\beta} (s_\theta)^\beta (1-s_\theta)^{M-\beta}. \tag{5.19}$$

The next proposition introduces a sufficient condition on the decision thresholds to get $\overline{p}^{\theta\theta} \to 1$ as $M \to \infty$.

Proposition 5.2. *If* $s_1 > s_2 > \ldots > s_K$ *and for all* $\theta \in \Theta \setminus \{K\}$*,*

$$\nu_\theta < s_\theta < \nu_{\theta-1},$$

then for all $\theta \in \Theta$*,*

$$\lim_{M\to\infty} \overline{p}^{\theta\theta} = 1, \qquad \theta \in \Theta. \tag{5.20}$$

Proof. See Appendix D.1. □

5.3.2 Case II

The results obtained in this case are in fact the extension of those shown in Section 4.5. Let $\overline{\mathbf{p}}^1 = (\overline{p}^{11}, \ldots, \overline{p}^{K1})$ and consider the functions

$$h_\theta(\overline{\mathbf{p}}^1) = \frac{\sum_{k\in\Theta} p_k \overline{p}^{k1} q(\theta,k)}{\sum_{k\in\Theta} p_k \overline{p}^{k1}}, \tag{5.21}$$

$$F_\theta(\overline{\mathbf{p}}^1) = \sum_{d=\lceil \nu_1 M \rceil}^{M} \binom{M}{\beta} \left(h_\theta(\overline{\mathbf{p}}^1)\right)^\beta \left(1 - h_\theta(\overline{\mathbf{p}}^1)\right)^{M-\beta}, \tag{5.22}$$

and $\mathbf{F}(\overline{\mathbf{p}}^1) = \left(F_1(\overline{\mathbf{p}}^1) \ldots F_K(\overline{\mathbf{p}}^1)\right)$. As an extension of Proposition 4.2, the following proposition provides a non-linear equation that has to be satisfied by $\overline{\mathbf{p}}^1$, to deduce the various $\overline{X}_\theta^{\tau,\beta}$ at equilibrium.

Proposition 5.3. *If the dynamic system (5.4) admits some equilibrium* $\overline{X}_\theta^{\tau,\beta}$*, then for any* $\theta \in \Theta$ *and* $\beta \leqslant \tau$*,*

$$\overline{X}_\theta^{\tau,\beta} = \begin{cases} 0, & \forall \tau < M, \\ \binom{M}{\beta} \left(h_\theta\left(\overline{\mathbf{p}}^1\right)\right)^\beta \left(1 - h_\theta\left(\overline{\mathbf{p}}^1\right)\right)^{M-\beta}, & \tau = M, \end{cases} \tag{5.23}$$

where $\overline{\mathbf{p}}^1$ *is the solution of*

$$\overline{\mathbf{p}}^1 = \mathbf{F}\left(\overline{\mathbf{p}}^1\right). \tag{5.24}$$

Again, using Brouwer's fixed-point theorem, the existence of a solution of (5.24) can be obtained, in Proposition 5.4.

Proposition 5.4. *For any* $0 = \nu_K < \cdots < \nu_1 < \nu_0 = 1$*, (5.24) always admits a solution, and the dynamical system (5.4) has an equilibrium.*

Proof. See Appendix D.2. □

Similar to Proposition 5.2, a sufficient condition to have $\lim_{M\to\infty} \overline{p}^{\theta\theta} = 1$ is stated in the following proposition.

Proposition 5.5. *If* $q(1,1) > q(2,1) \ldots > q(K,1)$ *and*

$$\max_{\theta\in\Theta\setminus\{1\}, k\in\Theta} q(\theta,k) < \nu_1 < q(1,1), \tag{5.25}$$

$$q(\theta+1,1) < \nu_\theta < q(\theta,1), \quad \forall \theta \in \Theta \setminus \{1\}, \tag{5.26}$$

then $\lim_{M\to\infty} \overline{p}^{\theta\theta} = 1$*,* $\forall \theta \in \Theta$*.*

Proof. See Appendix D.3. □

Explicit expressions for $\overline{p}^{\theta 1}$ are difficult to obtain from (5.24). Since $\overline{p}^{\theta 1}$ with $\theta \neq 1$ represent the proportions of agents that have wrongly estimated their group, the vector $\overline{\mathbf{p}}^1 = \left(\overline{p}^{11}, \ldots, \overline{p}^{K1}\right)$ should be close to $\widetilde{\mathbf{p}}^1 = \left(\widetilde{p}^{11}, \ldots, \widetilde{p}^{K1}\right)$. One has $\lim_{\mathbf{p}^1\to\widetilde{\mathbf{p}}^1} h_\theta\left(\mathbf{p}^1\right) = q(\theta,1)$. Assuming that at equilibrium, $h_\theta\left(\overline{\mathbf{p}}^1\right) \simeq q(\theta,1)$, using (5.23), $\overline{X}_\theta^{M,\beta}$ can be approximated as

$$\widetilde{X}_\theta^{M,\beta} = \binom{M}{\beta} \left(q(\theta,1)\right)^\beta \left(1 - q(\theta,1)\right)^{M-\beta}, \tag{5.27}$$

and follows thus a binomial distribution.

5.3.3 Choice of ν

Knowing $\overline{X}_\theta^{M,\beta}$ (in Case I) or $\widetilde{X}_\theta^{M,\beta}$ (in Case II), one is able to optimize the decision thresholds introduced in (5.2). The value of the ν_θs may for example be adjusted to maximize the correct decision rate (CDR) under some cumulated false decision rate (FDR) constraint evaluated using (5.27), but alternative performance requirements may be considered. The following proposition provides the optimal values of ν_θs to maximize the sum of CDR, *i.e.*, $\sum_{\theta\in\Theta} \overline{p}^{\theta\theta}$ or $\sum_{\theta\in\Theta} \widehat{p}^{\theta\theta}$.

Proposition 5.6. *In Case I, if*

$$\nu_\theta = \frac{\log\left(\frac{1-s_\theta}{1-s_{\theta+1}}\right)}{\log\left(\frac{s_{\theta+1}}{s_\theta}\frac{1-s_\theta}{1-s_{\theta+1}}\right)}, \qquad \forall\theta = 1\dots(K-1), \tag{5.28}$$

where s_θ is introduced in (5.16), then $\sum_{\theta\in\Theta}\overline{p}^{\theta\theta}$ takes its maximum value.

In Case II, if

$$\nu_\theta = \frac{\log\left(\frac{1-q(\theta,1)}{1-q(\theta+1,1)}\right)}{\log\left(\frac{q(\theta+1,1)}{q(\theta,1)}\frac{1-q(\theta,1)}{1-q(\theta+1,1)}\right)}, \qquad \forall\theta = 1\dots(K-1), \tag{5.29}$$

then $\sum_{\theta\in\Theta}\widetilde{p}^{\theta\theta}$ takes its maximum value.

Proof. See Appendix D.4. □

One sees that the values at equilibrium (5.18) and the decision thresholds in Case I are functions of the values of s_θ, themselves functions of the *a priori* proportions of agents in each group p_θ and of the characteristics of the LCT $q(\theta_i,\theta_j)$. In Case II, the approximate values at equilibrium (5.27) and the thresholds are only functions of the characteristics of the LCT. This may be very interesting in a practical implementation, since in Case II, agents do not need to know in advance the proportion of agents with a given status. Only the characteristics of the LCT are required.

5.4 Illustration

Consider a set $\mathcal{A}$ of N_A moving agents with very limited communication range, forming a DTN such that hypotheses A1-A4 are satisfied. Each agent is equipped with a sensor providing noisy observations

$$m_i(\boldsymbol{x}_i,t) = \phi(\boldsymbol{x}_i,t) + w_i(t), \qquad \forall i\in\mathcal{A} \tag{5.30}$$

of some scalar field $\phi(\boldsymbol{x}_i,t)$ at its location $\boldsymbol{x}_i$ and at time t. The components $w_i(t)$ of the measurement noise at time t in (5.30) described as realizations of independently distributed Gaussian variables $W_i\sim\mathcal{N}(e_i,\sigma^2)$, where e_i is some constant, agent-specific bias. The bias terms e_i, $i\in\mathcal{A}$, are assumed to be realizations of iid zero-mean Laplacian random variables E_i with parameter γ. Initially, the agents do not know their measurement bias.

The sensors with small e_i provide better measurements. Our aim is to apply the DSR algorithm introduced in Section 5.2 to allow each agent getting an estimate of its bias. For that purpose, introducing some constant parameter $\epsilon>0$, the set of agents is partitioned into K groups defined as

$$\mathcal{A}_\theta = \{i\in\mathcal{A} : \Lambda_{\theta-1}\leqslant|e_i|<\Lambda_\theta\},$$

where

$$\Lambda_\theta = \begin{cases}\epsilon\theta/\gamma, & \forall\theta\in\Theta\setminus\{K\},\\ \infty, & \theta=K.\end{cases}$$

Therefore, the proportion of agents with status θ is $p_\theta = \exp(-\epsilon(\theta-1)) - \exp(-\epsilon\theta)$. The status θ of an agent represents thus the group to which it belongs and provides an indication of the level of its bias. This problem is a generalization of that introduced in [Chi+11], where $\phi(\boldsymbol{x}_i, t)$ is assumed constant with location and time, and where the bias can only take discrete values.

Two agents, when meeting at some location $\boldsymbol{x}$ and time t, perform an independent measurement of $\phi(\boldsymbol{x}_i, t)$. Then, following Algorithm 5, both agents exchange their estimate of θ determine if they want to continue interacting. If this is the case, they have to exchange measurements and run individually the LCT introduced in Section 5.4.1.

5.4.1 LCT

The LCT considered in this work has similar form as the LODT considered in Chapters 3 and 4.

Consider a measurement m, some tolerance ω, and the interval $[m] = [m-\omega, m+\omega]$ of width 2ω centered around m. Consider now Agents i and j meeting at time t and exchanging the measurements m_i and m_j they just performed at the same location $\boldsymbol{x}$. The *set* estimate [BBC90] of $\phi(\boldsymbol{x}, t)$ obtained from m_i and m_j is defined as $\left[\widehat{\phi}(m_i, m_j)\right] = [m_i] \cap [m_j]$. If $\left[\widehat{\phi}(m_i, m_j)\right] \neq \emptyset$ it is likely that the biases e_i and e_j are close. Both agents can conclude that their sensors perform similarly. If $\left[\widehat{\phi}(m_i, m_j)\right] = \emptyset$, it is likely that e_i and e_j differ significantly more than ω. An agent is in this case unable to determine whether it is carrying the best sensor. As a consequence, each agent chooses to conclude that its sensor behaves worse than that of the other agent. One obtains as a consequence the following low-complexity LCT,

$$y_{i,j} = y_{j,i} = \begin{cases} 1, & \text{if } \left[\widehat{\phi}(m_i, m_j)\right] \neq \emptyset, \\ 0, & \text{else.} \end{cases} \tag{5.31}$$

For any arbitrary pair of status (θ, θ'), one is able to evaluate the probability $q(\theta, \theta') = \mathbb{P}\{Y_{i,j} = 1 | i \in \mathcal{A}_\theta, j \in \mathcal{A}_{\theta'}\}$ as a function of ω,σ, and γ.

5.4.2 Numerical verification of theoretical results

This section presents first the solution of the state equation (5.4) describing the evolution of the proportion of nodes in various states. Algorithm 5 is simulated first considering the Jumping model introduced in Section 4.7.1.

For the numerical example, one takes $K = 4$, $\sigma^2 = 0.16$, and $\gamma = \epsilon = 0.7$, resulting in

$$p_1 = 0.503,\ p_2 = 0.250,\ p_3 = 0.124,\ p_4 = 0.123. \tag{5.32}$$

Moreover, taking $\omega = 1.8$, one gets

$$\mathbf{q} = \begin{pmatrix} 0.95 & 0.66 & 0.22 & 0.01 \\ 0.66 & 0.52 & 0.44 & 0.10 \\ 0.22 & 0.44 & 0.50 & 0.27 \\ 0.01 & 0.10 & 0.27 & 0.35 \end{pmatrix}, \tag{5.33}$$

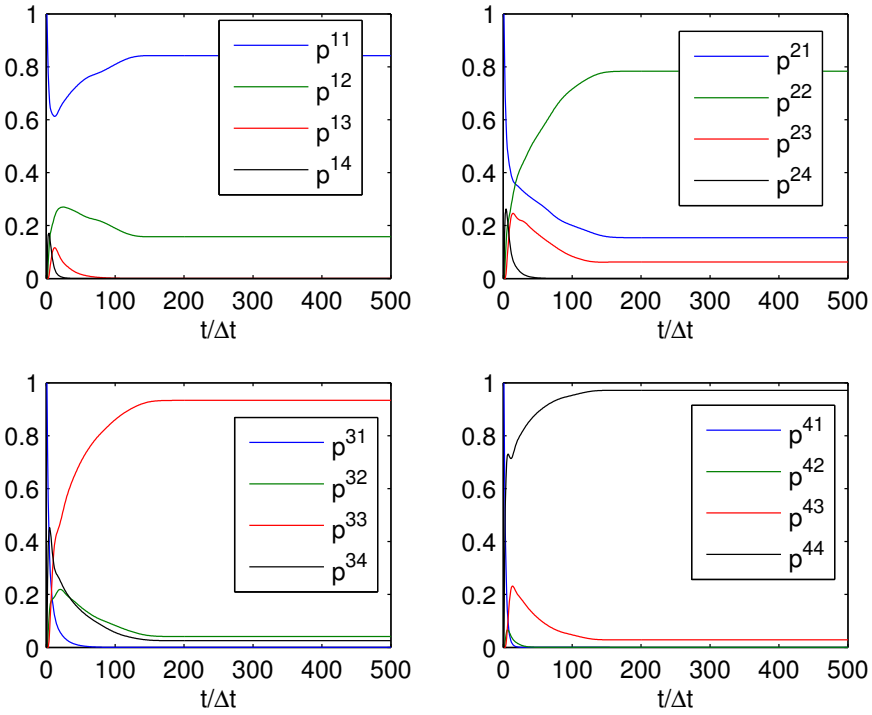

FIGURE 5.1: Case I: evolution of $p^{\theta\widehat{\theta}}(t)$ obtained solving (5.4).

which satisfies Assumption *A5*. Besides, one considers $M = 50$, and a sampling period Δt during which the inter-contact probability is $\lambda\Delta t = 0.33$. The decision thresholds are chosen using Proposition 5.6:

- In Case I, $\nu_1 = 0.61$, $\nu_2 = 0.42$, and $\nu_3 = 0.19$;
- In Case II, $\nu_1 = 0.84$, $\nu_2 = 0.43$, and $\nu_3 = 0.07$.

Figure 5.1 and 5.2 presents the evolution $p^{\theta\widehat{\theta}}(t)$ for $\theta, \widehat{\theta} \in \Theta = \{1, 2, 3, 4\}$, in Case I and in Case II respectively. One observes that the proportion agents of each state converges. Moreover, in Case II, for any $\theta \in \Theta$, $p^{\theta\theta}$ is close to 1 for t sufficient large, while $p^{\theta\widehat{\theta}}$ tends to 0 for any $\widehat{\theta} \neq \theta$. Nevertheless, in Case I, Algorithm 5 does not behave in a satisfying way: the values of p^{11} and p^{22} at the equilibrium are around 0.8, which means that 20% of the nodes with status $\theta = 1$ or $\theta = 2$ do not correctly rate their status.

In the following simulations, one considers the same Jumping model as presented in Section 4.7.1 that leads to $\lambda\Delta t \approx 0.33$. The evolutions of $p^{\theta\widehat{\theta}}(t)$ are shown in Figure 5.3 and 5.4 for Case I and Case II respectively. They are very close to those predicted by the direct integration of the state equation (5.4) as shown in Figure 5.1 and 5.2.

Figure 5.6 further illustrates the good match between theory and simulation for the proportions of states $\overline{X}_\theta^{50,\beta}$ at equilibrium in Case I. In Case II, Figure 5.6 shows also a very good match between the simulation and the approximation of $\overline{X}_\theta^{50,\beta}$ obtained using (5.27).

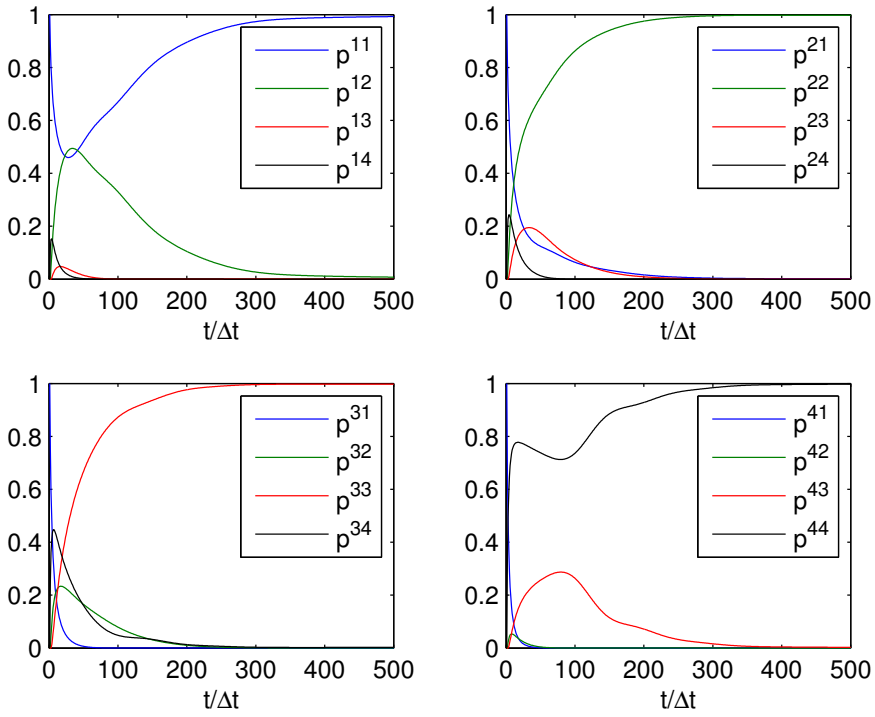

FIGURE 5.2: Case II: evolution of $p^{\theta\widehat{\theta}}(t)$ obtained solving (5.4).

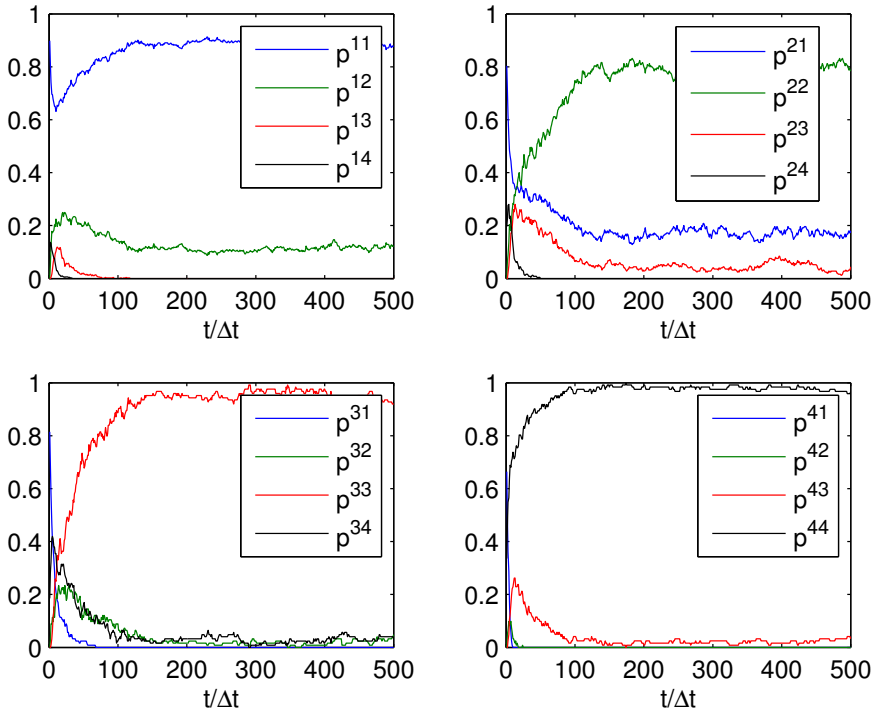

FIGURE 5.3: Case I: evolution of $p^{\theta\widehat{\theta}}(t)$ by simulations with the Jumping model.

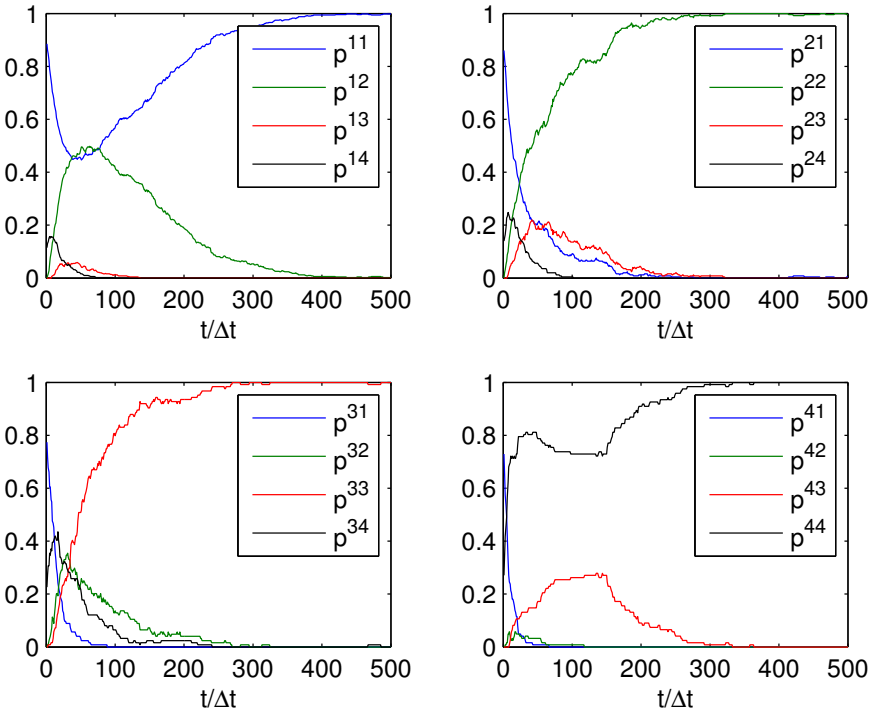

FIGURE 5.4: Case II: evolution of $p^{\theta\widehat{\theta}}(t)$ by simulations with the Jumping mode.

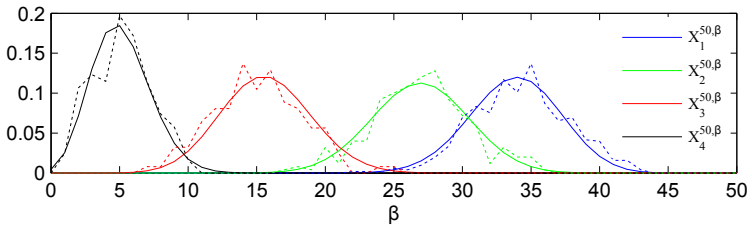

FIGURE 5.5: Case I: comparison of $\overline{X}_{\theta}^{50,d}$ at the equilibrium, lines are for the theoretical values obtained integrating (4.32) and dotted lines are for the moving agents simulation.

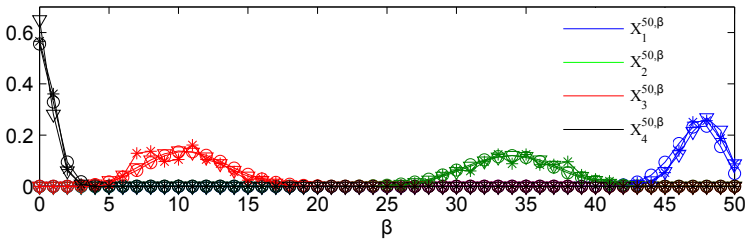

FIGURE 5.6: Case II: comparison of $\overline{X}_{\theta}^{50,d}$ at the equilibrium in, circles are for the theoretical values obtained integrating (5.4), triangles are the approximations obtained from (5.27), and crosses are for the moving agents simulation.

5.4.3 Simulation with real databases

In this section, one applies again the databases *Reality* and *Infocom05*. For each database, 100 independent simulations are performed and results are averaged over these simulations. As in Section 5.4.2, $K = 4$ groups are considered. The status of each node is randomly chosen according to (5.32). When an agent decides to continue interaction with an other agent, instead of performing a measurement and a test, a LCT is simulated with outcome randomly generated using (5.33).

Case I

At the top of Figure 5.7, for each time instant, a dot on the line corresponding to the index of an agent represents a contact with an other agent. The bottom of Figure 5.7 presents the evolution with time of the proportions $p_e^\theta = 1 - p^{\theta\theta}$ of nodes erroneously self-rating their status for each θ, considering $M = 50$ and $M = 200$. One also observes that the convergence speed of p_e^θ is highly related to the inter-contact rate (reflected by the density of points in the sub-figures at the top): variations are significant at beginning of working hours.

Figure 5.7 as well as the values at equilibrium of $\overline{X}_\theta^{M,\beta}$ reported in Figure 5.10 show some mismatch between theory and simulations for the agents with status $\theta = 1$ and $\theta = 2$. Although there are many inter-contacts, some $X_\theta^{M,\beta}$ s have not reached their theoretical values, especially for $\theta = 1$ and $\theta = 2$. This effect is more significant when M is small, which is consistent with Proposition 5.4: the behavior of the DSR algorithm improves when M increases. After investigation of the trace, one observes that some agents have only few contacts with other agents. For example, 4 agents have less than 100 contacts with other agents. Moreover, the variety of agents met by some agents is limited. For example, 4 nodes have contacts with less than 10 different other nodes. This has a significant influence on the performance of the algorithm.

Case II

Figure 5.9 shows the evolution of p_e^θ in Case II. For both databases, p_e^θ remains around 10^{-2} for all $\theta \in \Theta$ after a sufficient long time.

Figure 5.10 represents $\overline{X}_\theta^{M,\beta}$ obtained by using the databases *Reality* and *Infocom05*, and also by the approximation (5.10). In this case, there is an excellent match between the values at equilibrium predicted by theory and those obtained in simulation.

Comparison

Comparing the theoretical values of $\overline{X}_\theta^{50,\beta}$ at equilibrium for different values of $\theta \in \Theta$, one observes from Figures 5.8 and 5.10 that the different curves overlap more in Case I, especially those for $\theta = 1$ (blue) and $\theta = 2$ (green) and overlap much less in Case II. As a consequence, when considering some self-rating thresholds, better self-rating decision will be obtained in Case II. This illustrates that the importance of the probability of interaction $\alpha\left(\widehat{\theta}_i, \widehat{\theta}_j\right)$. In the considered application scenario, a better self-rating is obtained when agents only perform LCT with agents believing their status is the best one.

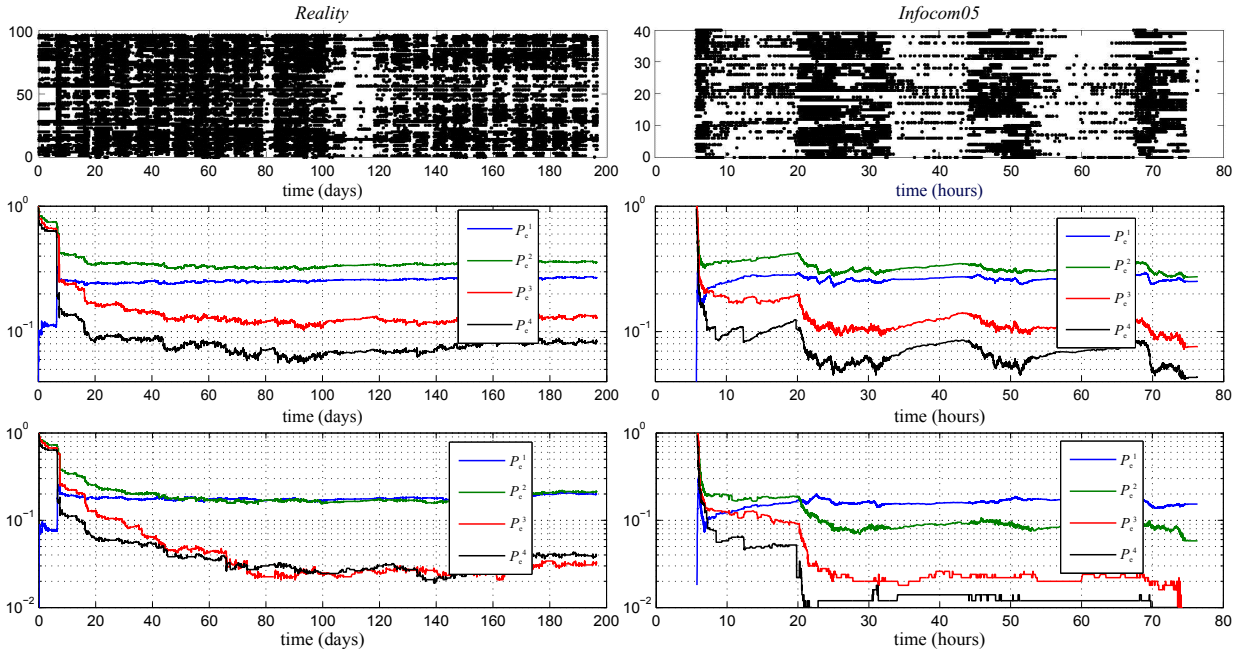

FIGURE 5.7: Case I: Indexes of active nodes (having met another node) at different time (top) and evolution of $P_e^\theta = 1 - p^{\theta\theta}$ when $M = 50$ (middle) and $M = 200$ (bottom) obtained using the *Reality* database (left) and the *Infocom05* database (right).

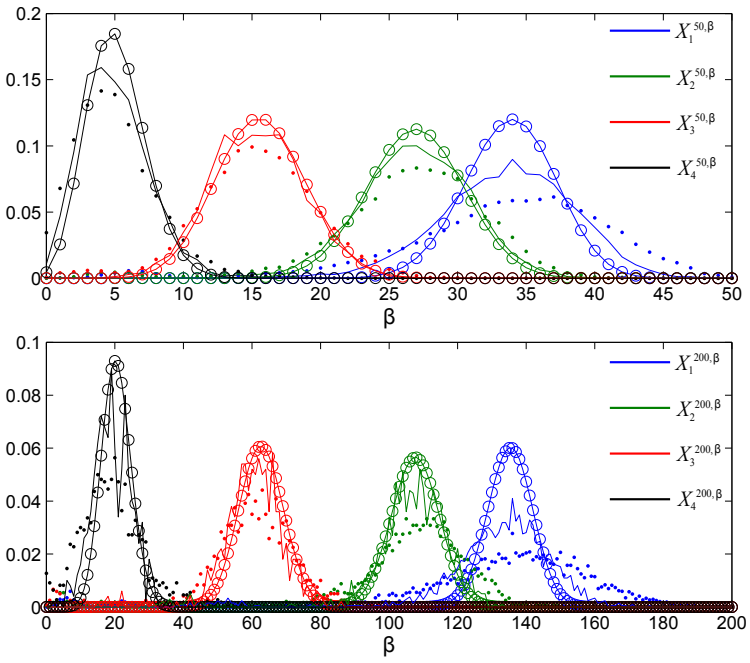

FIGURE 5.8: Case I: Values of $\overline{X}_\theta^{50,\beta}$ (top) and of $\overline{X}_\theta^{200,\beta}$ (bottom) at the end of the simulation: circles are for the theoretical values at equilibrium obtained from (5.4), dots are simulation results using the *Reality* database, and plain lines are simulation results using the *Infocom05* database.

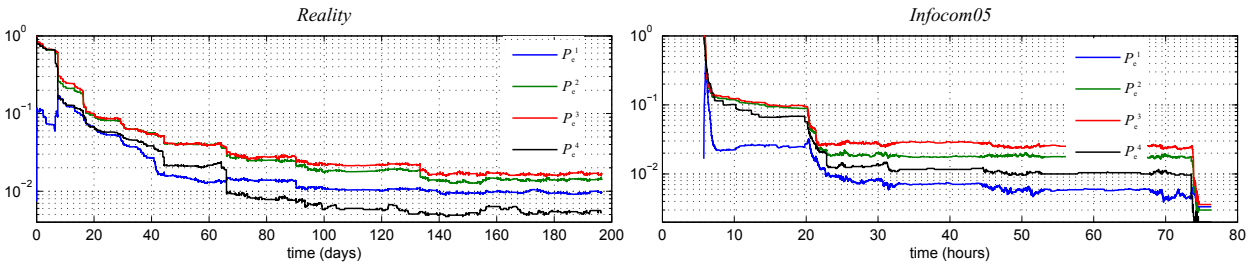

FIGURE 5.9: Case II: Evolution of $P_e^\theta = 1 - p^{\theta\theta}$ obtained using the *Reality* database (left) and the *Infocom05* database (right)

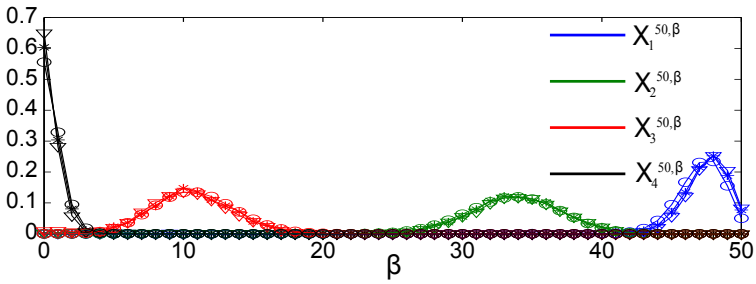

FIGURE 5.10: Case II: Values of $\overline{X}_\theta^{50,\beta}$ at the end of simulation: circles are for the theoretical values obtained from (5.4), triangles are simulation results using the *Reality* database, and crosses are simulation results using the *Infocom05* database.

5.5 Conclusions

This work has investigated the problem of helping agents self-rating their level of ability at doing some task via exchange of information with peers. Using local compatibility tests involving, *e.g.*, data exchanged during meetings with other agents, each agent is able to estimate the proportion of agents it is better at doing the considered task. With that information, each agent may then determine to which group of agents with similar expertise it belongs to.

The behavior of the proposed algorithm is described using dynamical equations. The existence of an equilibrium is established. The proportions of agents with similar beliefs in their expertise is characterized at equilibrium. This gives some insight in the tuning of the various parameters of the proposed algorithm.

The approach is illustrated with agents equipped with sensing devices of different sensing performance, which may be found in crowd sensing scenarios. Simulation results are in good match with theory.

Significant work remains to be done to analyze the behavior of the proposed algorithm with generic probabilities of interaction. The existence and uniqueness of the equilibrium has also to be shown in the general case. Nevertheless, the proposed approach may be useful to analyze other types of self-rating problems.

Chapter 6

Conclusion and Prospective

6.1 Conclusion

The first part of this thesis contains a detailed analysis of a correlated data gathering problem in WSNs. More precisely, the random linear network coding approach is considered to perform jointly data compression and data collection. A the components of compressible are observed by wireless sensors. An information-theoretic approach is applied to demonstrate the necessary and sufficient conditions to realize the asymptotically perfect reconstruction of this vector considering MAP estimation. The obtained results are consistent with those in the state-of-the-art and are more general, since a stationary and ergodic model of the vector components is considered as well as a sparse network-coding matrix. Moreover, both sensing noise and noise corrupting the network-coded components are taken into account. Interestingly, in absence of noise, the coding matrix can be arbitrarily sparse to achieve the optimum lower bound of the compression ratio. Some sufficient conditions on the minimum sparsity factor is proposed, which mainly depend on the additive noise.

The second part of this thesis concerns the distributed self-rating (DSR) problem, for wireless networks with nodes that have different expertise at performing some task (sensing, detection, classification...). The main assumption is that each node does not know its level of expertise and is willing to estimate it.

In Chapter 3, one considers a dense WSN and only two levels of expertise: network nodes are either equipped with normal sensors or with defective ones producing outliers. This thesis has proposed and analyzed the performance of two distributed faulty node detection (DFD) algorithms to help each node in determining whether it is equipped with a defective sensor. A node first collects data from its neighborhood, processes them to decide, using some generic local outlier detection test, whether these data contain outliers and broadcasts the result. Then, it determines the status (good or defective) of its own sensor using its result and those received from neighboring nodes. A single-decision and an iterative algorithm for DFD are proposed. Bounds on the performance of the single-decision algorithm are derived. A theoretical analysis of the probability of error and of the equilibrium of the iterative algorithm is provided for a wide class of local outlier detection tests. The trade-off between false alarm probability and detection probability is characterized theoretically and by simulation. MAC-layer issues, as well as the effect of packet losses are accounted for. The proposed algorithms have also been implemented on a real testbed to verify their efficiency.

In Chapter 4, the DFD problem is considered in delay tolerant networks (DTNs) which have sparse and dynamic connectivity. Propagation of faulty data in DTNs can be a critical aspect

to counteract due to the inherent feature of exhibiting frequent disconnections,. Indeed the rare meeting events require that nodes are effective and efficient in propagating correct information. Accordingly, mechanisms to rapidly identify possible faulty or misbehaving nodes should be obtained. DFD has been addressed in the literature in the context of dense WSNs, but seldom considered in DTNs. One proposes a fully distributed, easily implementable, and fast converging approach to allow each DTN node to rapidly identify whether its sensors are producing outliers. The difference between the DFD algorithms in Chapter 3 and that in Chapter 4 dedicated to DTNs mainly comes from the communication conditions, which also makes their analysis quite different. The behavior of the proposed algorithm is described by some continuous-time state equations, whose equilibrium is characterized. Interestingly, the proportions of nodes with different beliefs of their status at the equilibrium follow a Binomial distribution. Detection and false alarm rates are evaluated both theoretically and by simulation and show an excellent match. Theoretical results assess the effectiveness of the proposed solution and give guidelines in the design of the algorithm.

In Chapter 5, the DSR problem with agents with multiple levels of expertise is considered. The proposed self-rating algorithm involves pairwise interactions between agents and a local comparison test, able to determine which, among two agents, performs better. A probability of interaction has been introduced to determine whether two agents, interact during a meeting. The analysis of the dynamics of the DSR algorithm extends that considered in Chapter 4. Two special forms of the probability of interaction are analyzed in details. The main difficulty lies in determining decision thresholds to minimize the probability of erroneous self-rating. Simulation results match well theoretical results in the context of nodes equipped with sensors aiming at determining the performance of their sensors. These results have also shown the importance of an appropriate choice of the probability of interaction.

6.2 Perspective

Several open research directions may be considered at the end of this thesis.

6.2.1 Sparse random linear network coding

In Chapter 2, we have tried a first step towards handling varying levels of sparsity in the network coding matrix considering noisy environments to analyze data compression limits in WSNs.

To ensure the optimal compression ratio, some sufficient conditions on the sparsity of the network coding matrix have been exhibited. For instance, these results mainly depend on the minimum sparsity level, *i.e.*, $\min_{i=1\ldots N} \gamma_i$, recall that γ_i represents the sparsity factor of the i-th column of the network coding matrix. A future reserach direction should be related to conditions on the average sparsity level of the network coding matrix, *i.e.*, $\sum_{i=1}^{N} \gamma_i/N$. Besides, our converse analysis is mainly based on the classical Fano's inequality, thus no necessary condition on the sparsity of the network coding matrix has been proposed. Hence, it is also essential, yet challenging, to search some converse information-theoretic result, in order to see whether the necessary and sufficient conditions on the sparsity could converge.

Beside the results obtained related to the sparsity factor, another interesting problem is to see how to set up a WSN that meets these sparsity requirements. For example, when nodes are randomly and uniformly distributed over a plane with some density, one needs to determine with which probability mass function each node should choose the number of received packets to linearly combine, in order to satisfy the minimum sparsity requirement. It may be interesting to consider routing algorithms designed toadjust the sparsity of network coding matrix, see, *e.g.*, [Bas+12].

6.2.2 Distributed faulty node detection in a DTN

In Chapter 4, one has assumed that the status of the nodes remains constant, which facilitates the analysis of the DFD algorithm in a DTN. It would be interesting to consider a dynamic status of the nodes and see its impact. This problem is quite realistic. For example, a good node may become faulty when its power is low, a faulty node may become good when its defective sensor is replaced, or when it is able to harvest enough energy to reload its battery. In this situation, transition probabilities between status should be introduced and the Markov chains conditioned by the two status (*see, e.g.*, Figure 4.1, page 76) are no more independent.

The proposed approach to analyze the dynamics and equilibrium of the DFD algorithm may be useful to analyze other types of self-estimation problems. For example, the DFD algorithm may be applied in the context of malware detection. In this context, nodes may be infected by some malware, the latter propagating to other nodes during exchanges. This extension requires again that each node in the network is rational and willing to know whether it is infected by the malware. Two nodes which meet can interact to learn whether there is some suspicious actions during the interaction. Since the status of nodes changes with time, this clearly requires to consider transitions between node status. This makes the detection of nodes infected by malware more challenging. Specific malware propagation models (see [KSA12] for example) may be considered and introduced into our Markov model for the analysis of the dynamic. There will be a trade-off between the proportion of active infected nodes and the proportion of good nodes being wrongly detected as infected.

Finally, for the time being, a constrant decision threshold has been considered. A time-varying threshold may be of interest to speed-up the convergence of the algorithm to an equilibrium. This may be particularly useful in the context of malware detection. To perform such design, one cannot consider only dynamics at equilibrium. Transients have to be more carefully studied than was done in this thesis.

6.2.3 Distributed self-rating

As mentioned in Chapter 5, our results on distributed self-rating have been obtained considering two special types of interaction probabilities. To widen its application potential, significant work remains to be done to analyze the behavior of the proposed algorithm with generic probabilities of interaction. An interesting related problem is the optimization of the values ofthe interaction probabilities. Several criteria may be considered. One may try to maximize the proportion of correct self-rating Alternatively, considering some correct self-rating threshold, one may be interested at minimizing

the number of necessary interactions between agents. This is important when interactions have some cost, *e.g.*, are time consuming, drain the energy of agents...

This problem can be related to the well-known multi-armed bandits (MAB) problem [AB10], in which the focus is to learn the statistics of different machines to choose the best machine and maximize some reward. Here, instead of choosing at each time step the best arm to pull, in order to maximize, *e.g.*, the sum of discounted future rewards [SB98], the aim is to determine whether it may be of interest to interact with an encountered agent with a given estimate of its status. Interaction providing only little additional knowledge to the agent on its own status should be avoided. In a DTN, however, contrary to the classical MAB problem, agents are not totally free to chose the encountered agents. So it may take some time to meet the agents which will maximize this additional knowledge.

Another related problem is distributed ranking by pairwise comparison [Hec+16]. The distributed self-rating problem becomes a distributed ranking problem when the number of levels of ability is equal to the number of agents. The problem is relatively easy when some network controller is able to decide which pair of agents needs to interact at some time instant, in order to minimize the number of required tests. Nevertheless, in our distributed self-rating set-up, such controller does not exist. The only way to accelerate convergence of the algorithm is to adjust the interaction probability. For example, at the beginning of the algorithm, an interaction may always occur when two agents have a meeting. Then, after some time, the interaction probability becomes a function, that decreases with the difference of the estimated levels of the two agents. An attractive open problem is to learn the best compromise between the number of interaction and the convergence speed of the algorithm.

Appendix A

Appendices of Chapter 2

A.1 Proof of Lemma 2.1

Proof. Consider another matrix $\mathbf{B} \in \mathbb{F}_Q^{M\times N}$, whose entries are iid with the sparsity factor γ. According to [TBD12, Lemma 21], we have

$$\Pr\left\{\mathbf{B}_j\boldsymbol{\mu}^N = 0 \mid \left\|\boldsymbol{\mu}^N\right\|_0 = d_1\right\} = Q^{-1} + \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1}\left(1-Q^{-1}\right), \tag{A.1}$$

and

$$\Pr\left\{\mathbf{B}_j\boldsymbol{\mu}^N = q \mid \left\|\boldsymbol{\mu}^N\right\|_0 = d_1, q \in \mathbb{F}_Q \setminus \{0\}\right\} = Q^{-1} - \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1} Q^{-1}, \tag{A.2}$$

where $\mathbf{B}_j$ is j-th row of $\mathbf{B}$. If d_2 is the number of non-zero entries of $\mathbf{s}^M$, then by combining (A.1) and (A.2), one gets

$$\begin{aligned}
&\Pr\left\{\mathbf{B}\boldsymbol{\mu}^N = \mathbf{s}^M \mid \boldsymbol{\mu}^N \neq \mathbf{0}, \mathbf{s}^M\right\} = \prod_{j=1}^{M} \Pr\left\{\mathbf{b}_j\boldsymbol{\mu}^N = s_j \mid \boldsymbol{\mu}^N \neq \mathbf{0}, s_j\right\} \\
&= \left(Q^{-1} + \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1}\left(1-Q^{-1}\right)\right)^{M-d_2}\left(Q^{-1} - \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1} Q^{-1}\right)^{d_2} \\
&\leqslant \left(Q^{-1} + \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1}\left(1-Q^{-1}\right)\right)^{M}.
\end{aligned} \tag{A.3}$$

Now consider the matrix $\mathbf{A}$ with its columns having different sparsity factors such that $\gamma = \min_{i=1\ldots N}\gamma_i$. One has the following extension

$$\begin{aligned}
&\Pr\left\{\mathbf{A}\boldsymbol{\mu}^N = \mathbf{s}^M \mid \boldsymbol{\mu}^N \neq \mathbf{0}, \mathbf{s}^M\right\} = f\left(d_1, d_2\right) \leqslant f\left(d_1, 0\right) \\
&\leqslant \left(Q^{-1} + \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{d_1}\left(1-Q^{-1}\right)\right)^{M},
\end{aligned} \tag{A.4}$$

according to fact that the probability of having $\mathbf{A}\boldsymbol{\mu}^N = \mathbf{0}$ (with $\boldsymbol{\mu}^N \neq \mathbf{0}$)is larger if the matrix $\mathbf{A}$ is sparser. The monotonicity of the function is not hard to obtain with its expression and the condition (2.5). □

A.2 A Possible Situation for $\mathcal{H}(\Theta) = 0$

Consider N sensors uniformly deployed over a unit-radius disk. The physical quantities (in $\mathbb{R}$), which are collected by the sensors, are denoted by $\mathbf{\Omega}^N \in \mathbb{R}^N$. We assume that $\mathbf{\Omega}^N \sim \mathcal{N}(0, \mathbf{\Sigma})$ with

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \mathrm{e}^{-\lambda d_{1,2}^2} & \cdots & \mathrm{e}^{-\lambda d_{1,N}^2} \\ \mathrm{e}^{-\lambda d_{2,1}^2} & 1 & & \mathrm{e}^{-\lambda d_{2,N}^2} \\ \vdots & & \ddots & \vdots \\ \mathrm{e}^{-\lambda d_{N,1}^2} & \cdots & \cdots & 1 \end{bmatrix}, \tag{A.5}$$

where λ is some constant, $d_{i,j}$ is the distance between sensors i and j. The distance between two sensors is random since the location of each sensor is random. The real-valued entries of $\mathbf{\Omega}^N$ are quantized with a $Q-$level scalar quantizer. We assume that $Q = 2$, corresponding to the rule

$$\Theta_i = \begin{cases} 0 & \text{if } \Omega_i < 0, \\ 1 & \text{if } \Omega_i \geq 0. \end{cases} \tag{A.6}$$

With the above assumptions, we can prove the following lemma.

Lemma A.1. *The conditional entropy $H\left(\Theta_n \mid \mathbf{\Theta}_1^{n-1}\right)$ converges to zero for $n \to \infty$.*

Proof. Suppose that j is the index of the sensor which has the minimum distance to sensor n, among the $n-1$ neighbor sensors, *i.e.*,

$$j = \arg \min_{1 \leq i \leq n-1} d_{n,i}. \tag{A.7}$$

We have

$$H\left(\Theta_n \mid \mathbf{\Theta}_1^{n-1}\right) \leq H\left(\Theta_n \mid \Theta_j\right) \tag{A.8}$$

Denote the minimum distance as $\underline{d}(n) = d_{n,j}$, the covariance matrix of Ω_n and Ω_j is

$$\mathbf{\Sigma}_n = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \tag{A.9}$$

where $\rho = \mathrm{e}^{-\lambda \underline{d}(n)^2}$. For a pair of realizations ω_n and ω_j, the joint probability density function writes

$$g(\omega_n, \omega_j) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{\omega_n^2 + \omega_j^2 - 2\rho\omega_n\omega_j}{2(1-\rho^2)}\right). \tag{A.10}$$

We easily obtain the probability of both Ω_n and Ω_j being negative,

$$\begin{aligned} \Pr\{\Omega_n < 0 \text{ and } \Omega_j < 0\} &= \int_{-\infty}^{0}\int_{-\infty}^{0} g(\omega_n, \omega_j)\,\mathrm{d}\omega_n\mathrm{d}\omega_j \\ &= \frac{1}{4} + \frac{1}{2\pi}\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right) \end{aligned} \tag{A.11}$$

Taking into account (A.6), (A.11) is exactly the probability of the pair (Θ_n, Θ_j) being $(0,0)$. Define

$$\varepsilon(\rho) := \frac{1}{4} - \frac{1}{2\pi} \arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right). \tag{A.12}$$

After the similar derivations, one obtains

$$\Pr(\Theta_n = 0, \Theta_j = 0) = \Pr(\Theta_n = 1, \Theta_j = 1) = \frac{1}{2} - \varepsilon(\rho) \tag{A.13}$$

and

$$\Pr(\Theta_n = 0, \Theta_j = 1) = \Pr(\Theta_n = 1, \Theta_j = 0) = \varepsilon(\rho). \tag{A.14}$$

Then the joint entropy is

$$H(\Theta_n, \Theta_j) = 1 + H_2(2\varepsilon(\rho)). \tag{A.15}$$

Meanwhile $H(\Theta_j) = 1$, thanks to the 2-level uniform quantizer. Obviously

$$H(\Theta_n \mid \Theta_j) = H_2(2\varepsilon(\rho)) = H_2\left(2\varepsilon\left(\mathrm{e}^{-\lambda \underline{d}^2}\right)\right). \tag{A.16}$$

This conditional entropy is increasing in $\underline{d}$. When the number of sensors increases, the disk will be denser, and the minimum distance $\underline{d}$ goes smaller. Thus, $\underline{d}$ tends to 0 as $n \to \infty$, which implies that $H(\Theta_n \mid \Theta_j) \to 0$. According to (A.8), we conclude that $H\left(\Theta_n \mid \mathbf{\Theta}_1^{n-1}\right)$ also goes to zero as $n \to \infty$. □

Applying the chain rule, the entropy rate writes

$$\mathcal{H}(\Theta) = \lim_{N\to\infty} \frac{H(\Theta_1) + \sum_{n=2}^{N} H\left(\Theta_n \mid \mathbf{\Theta}_1^{n-1}\right)}{N}. \tag{A.17}$$

By Cesaro mean [CT06b, Theorem 4.2.3], $\mathcal{H}(\Theta) = 0$ as $H\left(\Theta_n \mid \mathbf{\Theta}_1^{n-1}\right) \to 0$.

Appendix B

Appendices of Chapter 3

B.1 Proof of Lemma 3.1

We start with the proof of (3.11).

Note that both q_{D} and q_{FA} represent the probability that the LODT yields 1. One has, for any $k \in \mathcal{A}$

$$\begin{aligned}\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right)=1\middle|\varphi\left(\mathcal{A}\right)=0\right\} &= \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right)=1, T\left(\mathbf{M}_{\mathcal{A}}\right)=0\middle|\varphi\left(\mathcal{A}\right)=0\right\} \\ &\quad + \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right)=1, T\left(\mathbf{M}_{\mathcal{A}}\right)=1\middle|\varphi\left(\mathcal{A}\right)=0\right\} \\ &\overset{(a)}{=} \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}\setminus\{k\}}\right)=1, T\left(\mathbf{M}_{\mathcal{A}}\right)=1\middle|\varphi\left(\mathcal{A}\right)=0\right\} \\ &\leqslant \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{A}}\right)=1\middle|\varphi\left(\mathcal{A}\right)=0\right\}, \end{aligned} \tag{B.1}$$

where (a) is obtained from Property 3.2. Combining (3.7) and (B.1), one gets

$$q_{\mathrm{FA}}\left(\mathbf{M}_{\mathcal{A}\setminus\{K\}}\right) \leqslant q_{\mathrm{FA}}\left(\mathbf{M}_{\mathcal{A}}\right). \tag{B.2}$$

Suppose that $|\mathcal{A}| = n_{\mathrm{g}}+1$ with $n_{\mathrm{g}} \geqslant 1$, then based on Property 3.1, (B.2) is equivalent to $q_{\mathrm{FA}}\left(n_{\mathrm{g}}\right) \leqslant q_{\mathrm{FA}}\left(n_{\mathrm{g}}+1\right)$, which proves (3.11).

One may prove (3.12) and (3.13) in a similar way, by considering $\varphi\left(\mathcal{A}\right)=1$ in the derivations.

B.2 Analysis of the decision rule (3.22)

At the end of Algorithm 1, each node estimates its status from the results of the LODTs, which may be gathered in the vector

$$\boldsymbol{Y}_i^{(L)} = \left[Y_j^{(\ell)}\right]_{j\in\mathcal{B}_i^{(\ell)}, 1\leq\ell\leq L}. \tag{B.3}$$

Let $\boldsymbol{y}_i^{(L)}$ be one realization of $\boldsymbol{Y}_i^{(L)}$. In what follows, one will show that the decision rule (3.22) corresponds, under some simplifying assumptions, to

$$\widehat{\theta}_i = \arg\max_{\phi\in\{0,1\}} c_\phi \mathbb{P}\left\{\boldsymbol{Y}_i^{(L)} = \boldsymbol{y}_i^{(L)} \mid \theta_i = \phi\right\}, \tag{B.4}$$

where c_ϕ is some weight. If $c_0 = c_1 = 1$, then (B.4) is a maximum likelihood estimate of θ_i from $\boldsymbol{y}_i^{(L)}$; if $c_\phi = \mathbb{P}\left\{\theta_i = \phi\right\}$, then (B.4) is the maximum *a posteriori* estimate of θ_i.

Expressing the likelihoods in (B.4) is complicated due to the correlation between the components of $\boldsymbol{Y}_j^{(L)}$. Nevertheless, assuming that the $Y_j^{(\ell)}$s are iid, one obtains

$$\mathbb{P}\left\{\boldsymbol{Y}_i^{(L)} = \boldsymbol{y}_i^{(L)} \mid \theta_i = \phi\right\} = \prod_{\ell=1}^{L} \prod_{j \in \mathcal{B}_i^{(\ell)}} \mathbb{P}\left\{Y_j^{(\ell)} = y_j^{(\ell)} \mid \theta_i = \phi\right\}. \tag{B.5}$$

Assuming further that $\mathbb{P}\left\{Y_j^{(\ell)} = y_j^{(\ell)} \mid \theta_i = \phi\right\} = \mathbb{P}\left\{Y_i^{(\ell)} = y_i^{(\ell)} \mid \theta_i = \phi\right\}$ for all $j \in \mathcal{B}_i^{(\ell)}$, which can be justified by Property 3.1 when the number of measurements considered by Node j to get $Y_j^{(\ell)}$ is equal to the number of measurements by Node i to get $Y_i^{(\ell)}$. Then for all $j \in \mathcal{B}_i^{(\ell)}$, one gets

$$\mathbb{P}\left\{Y_j^{(\ell)} = 1 \mid \theta_i = 0\right\} = f_1\left(\bar{\mu}_{\mathrm{g}}, \bar{\mu}_{\mathrm{d}}\right), \tag{B.6}$$

$$\mathbb{P}\left\{Y_j^{(\ell)} = 1 \mid \theta_i = 1\right\} = f_2\left(\bar{\mu}_{\mathrm{g}}, \bar{\mu}_{\mathrm{d}}\right). \tag{B.7}$$

Let $z_i = \sum_{\ell=1}^{L} \sum_{j \in \mathcal{B}_i^{(\ell)}} y_j^{(\ell)} = z_i$ and $n_i = \sum_{\ell=1}^{L} \left|\mathcal{B}_i^{(\ell)}\right|$, then from (B.5-B.6) one obtains

$$\frac{c_0 \mathbb{P}\left\{\boldsymbol{Y}_i^{(L)} = \boldsymbol{y}_i^{(L)} \mid \theta_i = 0\right\}}{c_1 \mathbb{P}\left\{\boldsymbol{Y}_i^{(L)} = \boldsymbol{y}_i^{(L)} \mid \theta_i = 1\right\}} = \frac{c_0 f_1^{z_i} \left(1 - f_1\right)^{n_i - z_i}}{c_1 f_2^{z_i} \left(1 - f_2\right)^{n_i - z_i}}. \tag{B.8}$$

Therefore, one obtains the following decision rule

$$\widehat{\theta}_i = \begin{cases} 0, & \text{if } \frac{z_i}{n_i} < \frac{\frac{1}{n_i} \log \frac{c_1}{c_0} - \log \frac{1-f_1}{1-f_2}}{\log \frac{f_1(1-f_2)}{f_2(1-f_1)}}, \\ 1, & \text{otherwise.} \end{cases} \tag{B.9}$$

When $c_0 = c_1 = 1$, the threshold corresponding to the maximum likelihood estimate is

$$\gamma_{\mathrm{ML}} = \frac{\log \frac{1-f_1}{1-f_2}}{\log \frac{1-f_1}{1-f_2} + \log \frac{f_2}{f_1}}. \tag{B.10}$$

When $c_\phi = \mathbb{P}\left\{\theta_i = \phi\right\}$, the threshold corresponding to the maximum *a posteriori* estimate of θ depends on n_i as follows

$$\gamma_{\mathrm{MAP}}\left(n_i\right) = \frac{\log \frac{1-f_1}{1-f_2} - \frac{1}{n_i} \log \frac{\bar{\mu}_{\mathrm{d}}}{\bar{\mu}_{\mathrm{g}}}}{\log \frac{1-f_1}{1-f_2} + \log \frac{f_2}{f_1}}. \tag{B.11}$$

One notes that $\lim_{n_i \to \infty} \gamma_{\mathrm{MAP}}\left(n_i\right) = \gamma_{\mathrm{ML}}$.

Figure B.1 represents the theoretical values of γ_{ML}, $\gamma_{\mathrm{MAP}}\left(\bar{\mu}\right)$ with $L = 1$, and $\gamma_{\mathrm{MAP}}\left(\bar{\mu}\right)$ with $l = 10$ as functions of $\bar{\mu}$ for different values of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$. For a constant value of $\overline{\mu}$, γ increases with $\bar{\mu}_{\mathrm{d}}$. Moreover, γ_{ML} and $\gamma_{\mathrm{MAP}}\left(\bar{\mu}\right)$ with $L = 10$ behave very similarly, increasing with $\bar{\mu}$. Finally, $\gamma_{\mathrm{MAP}}\left(\bar{\mu}\right)$ with $L = 1$ is always larger than the other thresholds: when a limited number of measurements is available, this limits the probability of false alarm.

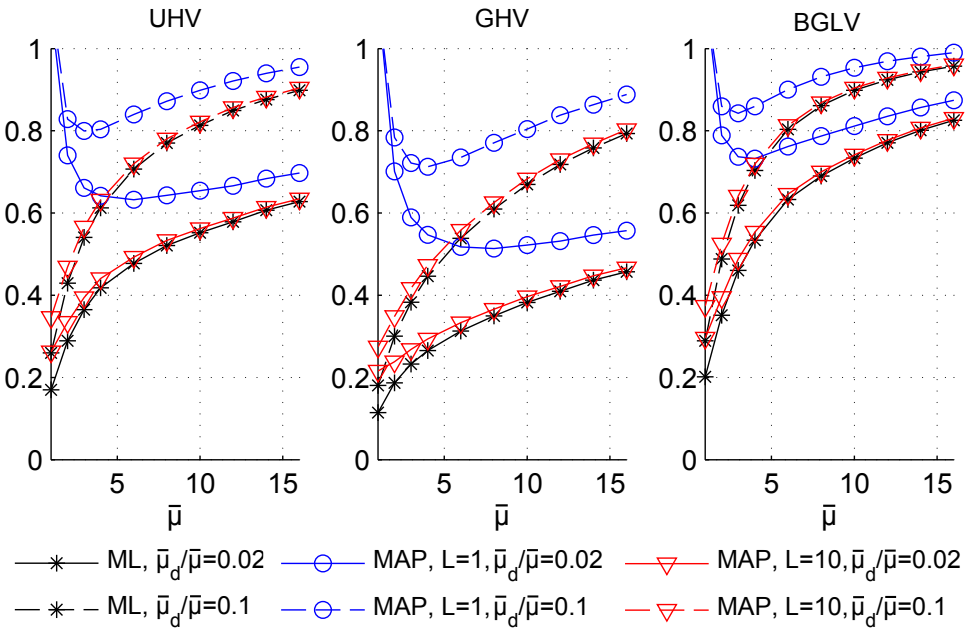

FIGURE B.1: γ_{ML}, $\gamma_{\mathrm{MAP}}(\bar{\mu})$ with $L=1$, and $\gamma_{\mathrm{MAP}}(\bar{\mu})$ with $L=10$ as functions of $\bar{\mu}$ for different values of $\bar{\mu}_{\mathrm{d}}/\bar{\mu}$.

B.3 Multi-hop Algorithm

When the WSN is sparse, *i.e.*, when $\bar{\mu}$ is small, the performance of DFD algorithms may be poor, due to a reduced amount of data to perform LODT. This appendix presents the modifications to be performed in the DFD algorithm to allow multi-hop data collection and LODT result dissemination. It is a variant of multi-hop dissemination protocol, see, *e.g.*, [Ver+10, Chap. 4.5], where data aggregation in packets is performed to limit the number of transmitted packets.

This feature requires that each node acts as a router to forward the packets it has received. This forwarding has to be limited using some time-to-live information T^{TL} associated to each data. T^{TL} has to be initialized to H, the maximum number of allowed hops. Let $\mathbf{p}_i^{(\ell,\mathrm{I},k)}$ be the payload of the k-th packet transmitted by Node i, during Phase I in the ℓ-th round. Initially $k=1$ and one has

$$\mathbf{p}_i^{(\ell,\mathrm{I},1)} = \left\{ \left(\begin{array}{cc} T_i^{\mathrm{TL}}, & m_i^{(\ell)} \end{array} \right) \right\}, \tag{B.12}$$

with $T_i^{\mathrm{TL}} = H$. The identification of Node i, stored in Addr_i, is included in the packet header. Node i has received packets from its neighbors in the set $\mathcal{N}_i^{(\ell,\mathrm{I},1)}$ during the first transmission of Phase I of Round ℓ. The corresponding data and associated T^{TL} are denoted $\left(\begin{array}{cc} T_j^{\mathrm{TL}} & , m_j^{(\ell)} \end{array} \right)$, $j \in \mathcal{N}_i^{(\ell,\mathrm{I},1)} \subset \mathcal{S}$. Node i then needs to forward all data which have a ${T_j}^{\mathrm{TL}} > 0$. To save bandwidth resources, these data are aggregated into one packet to get

$$\mathbf{p}_i^{(\ell,\mathrm{I},2)} = \bigcup_{j \in \mathcal{N}_i^{(\ell,\mathrm{I},1)} \text{ st } T^{\mathrm{TL}}>0} \left\{ \left(\mathrm{Addr}_j, \ T_j^{\mathrm{TL}} - 1, \ \ m_j^{(\ell)} \right) \right\}. \tag{B.13}$$

During k-th transmission with $k \geq 2$, packets are transmitted similarly, accounting for newly received data.

A similar process is performed during Phase II where LODT results are broadcast with limited hop count. Now, the payload of the first packet transmitted by Node i, during Phase II in the ℓ-th round is

$$\mathbf{p}_i^{(\ell,\mathrm{II},1)} = \left[\; T_i^{\mathrm{TL}}, \quad y_i^{(\ell)} \quad , \left|\mathcal{V}_i^{(\ell,\mathrm{I})}\right|, \quad \bigcup\nolimits_{j\in\mathcal{V}_i^{(\ell,\mathrm{I})}} \{\mathrm{Addr}_j\} \; \right]. \tag{B.14}$$

In (B.14), T_i^{TL} is again initialized at H. The number of nodes participating to the LODT results $y_i^{(\ell)}$ as well as their addresses are also indicated. For the k-th transmitted packets, payload containing LODT outcomes with $T^{\mathrm{TL}} > 0$ can again be aggregated as done in (B.13).

B.4 Proof of (3.52)

Consider a bounded sequence $0 \leqslant B(n) \leqslant 1$ for any $n \in \mathbb{N}$ and a real-valued function $G_N(\mu) = \sum_{n=0}^{N} B(n) \cdot \frac{\mu^n}{n!} \exp(-\mu)$ with $\mu \geqslant 0$. First, one evaluates the derivative of $G_N(\mu)$,

$$\frac{\mathrm{d}G_N(\mu)}{\mathrm{d}\mu} = \exp(-\mu) \cdot \left(\sum_{n=0}^{N} \left(B(n+1) - B(n)\right) \frac{\mu^n}{n!} - B(N+1) \frac{\mu^N}{N!} \right). \tag{B.15}$$

Second, one shows that $\frac{\mathrm{d}G_N(\mu)}{\mathrm{d}\mu}$ converges uniformly to

$$H(\mu) = \sum_{n=0}^{\infty} \left(B(n+1) - B(n)\right) \frac{\mu^n}{n!} \exp(-\mu). \tag{B.16}$$

Let $\varepsilon \in \mathbb{R}^+$ be an arbitrary number, then for any $\mu \geqslant 0$, there exists $N_\mu \in \mathbb{N}^+$, such that

$$\sum_{n=N_\mu}^{\infty} \frac{\mu^n}{n!} \exp(-\mu) < \varepsilon. \tag{B.17}$$

One has

$$\begin{aligned}
&\left| \frac{\mathrm{d}G_{N_\mu}(\mu)}{\mathrm{d}\mu} - H(\mu) \right| \\
&\leqslant B(N_\mu + 1) \cdot \frac{\mu^{N_\mu}}{N_\mu!} \exp(-\mu) + \sum_{n=N_\mu+1}^{\infty} \left|B(n+1) - B(n)\right| \cdot \frac{\mu^n}{n!} \exp(-\mu) \\
&\overset{(a)}{\leqslant} \sum_{n=N_\mu}^{\infty} \frac{\mu^n}{n!} \exp(-\mu) \overset{(b)}{<} \varepsilon,
\end{aligned}$$

where (a) comes from the fact that $0 \leqslant B(n) \leqslant 1$ and $-1 \leqslant B(n+1) - B(n) \leqslant 1$ for any $n \in \mathbb{N}$, and (b) is by (B.17). Similarly, one can also show that $G_N(\mu)$ converges uniformly to $G(\mu) = \sum_{n=0}^{N} B(n) \cdot \frac{\mu^n}{n!} \exp(-\mu)$ for any $\mu \geqslant 0$.

Therefore, according to [Rud64, Thm 7.17], one obtains that $\lim_{N\to\infty} \frac{\mathrm{d}G_N(\mu)}{\mathrm{d}\mu} = H(\mu)$, $G(\mu)$ is differentiable for any $\mu \geqslant 0$ and $\frac{\mathrm{d}G(\mu)}{\mathrm{d}\mu} = H(\mu)$.

B.5 Proof of Lemma 3.3

Let $\mathcal{R}_i = \{j \in \mathcal{S} \text{ such that } r_{i,j} \leqslant 2R_0\}$ be the set of nodes at a distance to the reference Node i less than $2R_0$, including i itself. For any $j \in \mathcal{B}_i$, $\mathcal{V}_j^{(\mathrm{I})} \subseteq \mathcal{R}_i$. Let $\mathbf{M}_{\mathcal{R}_i}$ and $\mathbf{M}_{\mathcal{V}_j^{(\mathrm{I})}}$ be the vectors of data provided by the sensors in $\mathcal{R}_i$ and $\mathcal{V}_j^{(\mathrm{I})}$ respectively. From Property 3.2, one has

$$\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_j^{(\mathrm{I})}}\right) = 0 \mid T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} = 1, \ \forall j \in \mathcal{B}_i. \tag{B.18}$$

To lighten notations, define $Z_i = \sum_{j\in\mathcal{B}_i} T\left(\mathbf{M}_{\mathcal{V}_j^{(\mathrm{I})}}\right)$. From (B.18), one deduces

$$\mathbb{P}\left\{Z_i = 0 \mid T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} = 1. \tag{B.19}$$

Then

$$\begin{aligned}
\mathbb{P}\left\{Z_i = 0\right\} &= \sum_{y=0}^{1} \mathbb{P}\left\{Z_i = 0, T\left(\mathbf{M}_{\mathcal{R}_i}\right) = y\right\} \\
&\geqslant \mathbb{P}\left\{Z_i = 0, T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} \\
&= \mathbb{P}\left\{Z_i = 0 \mid T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} \cdot \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} \\
&= \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\}.
\end{aligned} \tag{B.20}$$

On the other hand

$$\mathbb{P}\left\{Z_i < \left(N_{\mathrm{g}} + N_{\mathrm{d}}\right)\gamma\right\} \geqslant \mathbb{P}\left\{Z_i = 0\right\}. \tag{B.21}$$

Combining (B.20) and (B.21), one gets

$$\mathbb{P}\left\{Z_i \geqslant \left(N_{\mathrm{g}} + N_{\mathrm{d}}\right)\gamma\right\} \leqslant 1 - \mathbb{P}\left\{Z_i = 0\right\} \leqslant \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 1\right\}, \tag{B.22}$$

independently of the status of Node i. When $\theta_i = 0$,

$$\begin{aligned}
P_{\mathrm{FA}}\left(\mu_{\mathrm{g}}^{(\mathrm{II})}, \mu_{\mathrm{d}}^{(\mathrm{II})}\right) &\leqslant \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 1 \mid \theta_i = 0\right\} \\
&= \sum_{n_{\mathrm{g}}=1}^{\infty} \sum_{n_{\mathrm{d}}=0}^{\infty} h\left(n_{\mathrm{g}}, n_{\mathrm{d}}\right) \frac{\left(4\bar{\mu}_{\mathrm{d}}\right)^{n_{\mathrm{d}}-1}\left(4\bar{\mu}_{\mathrm{g}}\right)^{n_{\mathrm{g}}}}{\left(n_{\mathrm{g}}-1\right)! n_{\mathrm{d}}!} \exp\left(-4\bar{\mu}_{\mathrm{d}} - 4\bar{\mu}_{\mathrm{g}}\right),
\end{aligned}$$

leading to (3.36). Note that the area of the disk associated to $\mathcal{R}_i$ is $4\pi R_0^2$, thus in average $|\mathcal{R}_i \cap \mathcal{G}|$ is $4\bar{\mu}_{\mathrm{g}}$ and in average $|\mathcal{R}_i \cap \mathcal{D}|$ is $4\bar{\mu}_{\mathrm{d}}$. One may show (3.37) in a similar way.

B.6 Proof of Lemma 3.6

To prove that $g_{\mathrm{FA}}(\mu_{00},\mu_{10})$ as defined in (3.55) is monotone increasing in μ_{00}, one has to show that

$$\frac{\partial g_{\mathrm{FA}}}{\partial \mu_{00}} = \frac{\mu_{00}}{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}-\mu_{00}} \sum_{n_{\mathrm{g}}=1}^{\infty}\sum_{n_{\mathrm{d}}=0}^{\infty} \frac{\mu_{10}^{n_{\mathrm{d}}}\mu_{00}^{n_{\mathrm{g}}-1}\exp\left(-\mu_{10}-\mu_{00}\right)}{n_{\mathrm{d}}!\left(n_{\mathrm{g}}-1\right)!} \cdot\left(\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right)+\frac{\mu_{00}^2-\alpha_3\zeta\mu_{00}\bar{\mu}_{\mathrm{g}}+\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}{\mu_{00}\left(\alpha_3\zeta\bar{\mu}_{\mathrm{g}}-\mu_{00}\right)}\tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)\right) \tag{B.23}$$

is strictly positive. Therefore, a sufficient condition to have $\frac{\partial g_{\mathrm{FA}}}{\partial \mu_{00}} > 0$ is

$$\frac{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right)}{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)} > -\frac{\mu_{00}^2-\alpha_3\zeta\mu_{00}\bar{\mu}_{\mathrm{g}}+\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}{\left(\alpha_3\zeta\bar{\mu}_{\mathrm{g}}-\mu_{00}\right)\mu_{00}}. \tag{B.24}$$

Since

$$-\frac{\mu_{00}^2-\alpha_3\zeta\mu_{00}\bar{\mu}_{\mathrm{g}}+\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}{\left(\alpha_3\zeta\bar{\mu}_{\mathrm{g}}-\mu_{00}\right)\mu_{00}} = 1-\frac{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}{\left(\alpha_3\zeta\bar{\mu}_{\mathrm{g}}-\mu_{00}\right)\mu_{00}} \leqslant 1-\frac{4}{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}},$$

a sufficient condition to have (B.24) is

$$\frac{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right)}{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)} > \left(1-\frac{4}{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}\right)$$

for any $n_{\mathrm{g}} \geqslant 1$ and $n_{\mathrm{d}} \geqslant 0$, which is equivalent to

$$\min_{n_{\mathrm{g}}\geqslant 1, n_{\mathrm{d}}\geqslant 0} \frac{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right)}{\tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)} > 1-\frac{4}{\alpha_3\zeta\bar{\mu}_{\mathrm{g}}}. \tag{B.25}$$

One needs to find a lower bound of $\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right)$ as a function of $\tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)$. Introducing $\mathcal{I}=\{\theta_i=0, N_{\mathrm{g}}=n_{\mathrm{g}}+1, N_{\mathrm{d}}=n_{\mathrm{d}}\}$, by definition,

$$\begin{aligned}
\tau_{\mathrm{FA}}\left(n_{\mathrm{g}}+1,n_{\mathrm{d}}\right) &= \mathbb{P}\left\{\frac{\sum_{j\in\mathcal{B}_i}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}+1}=1\middle|\mathcal{I}\right\} \\
&\overset{(a)}{=} \mathbb{P}\left\{\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1\middle|\theta_K=0,\mathcal{I}\right\}\mathbb{P}\left\{Y_K=1\middle|\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1,\theta_K=0,\mathcal{I}\right\} \\
&\overset{(b)}{\geqslant} \mathbb{P}\left\{\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y'_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1\middle|\theta_K=0,\mathcal{I}\right\}\mathbb{P}\left\{Y_K=1\middle|\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1,\theta_K=0,\mathcal{I}\right\} \\
&\overset{(c)}{=} \mathbb{P}\left\{\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y'_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1\middle|\theta_i=\theta_K=0,N'_{\mathrm{g}}=n_{\mathrm{g}},N_{\mathrm{d}}=n_{\mathrm{d}}\right\} \\
&\qquad\cdot\mathbb{P}\left\{Y_K=1\middle|\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1,\theta_K=0,\mathcal{I}\right\} \\
&\overset{(d)}{=} \tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)\cdot\mathbb{P}\left\{Y_K=1\middle|\frac{\sum_{j\in\mathcal{B}_i\setminus\{K\}}Y_j}{n_{\mathrm{g}}+n_{\mathrm{d}}}=1,\theta_K=0,\mathcal{I}\right\} \\
&\overset{(e)}{\geqslant} \tau_{\mathrm{FA}}\left(n_{\mathrm{g}},n_{\mathrm{d}}\right)\cdot\mathbb{P}\left\{Y_K=1|Y_i=1,\theta_K=0,\mathcal{I}\right\},
\end{aligned} \tag{B.26}$$

where the considered $\mathcal{B}_i$ is such that $\theta_i = 0$, $N_{\mathrm{g}} = n_{\mathrm{g}} + 1$, and $N_{\mathrm{d}} = n_{\mathrm{d}}$. For all $j \in \mathcal{B}_i$, the LODTs are performed based on the data vector $\mathbf{M}_{\mathcal{V}_j^{(\mathrm{I})}}$ with outcome Y_j. In (a), $K \neq i$ is a random node in $\mathcal{N}_i \cap \mathcal{G}$; such node exists since $n_{\mathrm{g}} \geqslant 1$. In (b), $Y_j' = T\left(\mathbf{M}_{\mathcal{V}_j^{(\mathrm{I})} \setminus \{K\}}\right)$ and one uses the results of Lemma 3.1 to get $\mathbb{P}\{Y_j = 1\} \geqslant \mathbb{P}\left\{Y_j' = 1\right\}$ for all $j \in \mathcal{B}_i \setminus \{K\}$. To get (c), one uses the fact that Node K, equipped with a good sensor, is not used to get Y_j' and one accounts only for the presence of $N_{\mathrm{g}}' = n_{\mathrm{g}}$ nodes with good sensors. Then (d) is by definition of $\tau_{\mathrm{FA}}(n_{\mathrm{g}}, n_{\mathrm{d}})$ and see Appendix B.7 for the proof of (e). According to (B.25) and (B.26), the first statement of Lemma 3.6 is proved.

In the similar way, a sufficient condition to have $\mu_{10} P_{\mathrm{D}}(\mu_{00}, \mu_{10}) / (\bar{\mu}_{\mathrm{d}} - \mu_{10})$ an increasing function of μ_{10} is that

$$\begin{aligned} \frac{\tau_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}}+1)}{\tau_{\mathrm{D}}(n_{\mathrm{g}}, n_{\mathrm{d}})} &\geqslant \mathbb{P}\{Y_k = 1 \mid \theta_i = \theta_k = 1\} \\ &\geqslant q_{\mathrm{D}}(0,2) > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\mathrm{d}}}, \end{aligned} \tag{B.27}$$

which corresponds to the second statement of Lemma 3.6.

B.7 Proof of $\mathbb{P}\{Y_{j_1}=1\} \leqslant \mathbb{P}\{Y_{j_1}=1 \mid Y_i=1, Y_{j_2}=1, \dots\}$

In this section, one aims to prove the following lemma.

Lemma B.1. *Consider a LODT satisfying Properties 3.1 and 3.2, and some Node i. For any set of distinct indexes $\{i, j_1, j_2, \dots\} \subset \mathcal{B}_i$, one has*

$$\mathbb{P}\{Y_{j_1}=1\} \leqslant \mathbb{P}\{Y_{j_1}=1 \mid Y_i=1\} \leqslant \mathbb{P}\{Y_{j_1}=1 \mid Y_i=1, Y_{j_2}=1\} \leqslant \dots. \tag{B.28}$$

Proof. Let $Y_{ik} = T\left(\mathbf{M}_{\mathcal{V}_i^{(\mathrm{I})} \cap \mathcal{V}_k^{(\mathrm{I})}}\right)$. Note that if $\left|\mathcal{V}_i^{(\mathrm{I})} \cap \mathcal{V}_k^{(\mathrm{I})}\right| \leqslant 1$, one has $Y_{ik} = 0$.

$$\begin{aligned} &\mathbb{P}\{Y_k = 1 \mid Y_i = 1\} \\ &= \sum_{y=0}^{1} \mathbb{P}\{Y_k = 1, Y_{ik} = y \mid Y_i = 1\} \\ &\overset{(a)}{=} \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} + \mathbb{P}\{Y_k = 1, Y_{ik} = 0 \mid Y_i = 1\} \\ &\overset{(b)}{=} \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} + \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\} \\ &= \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} (1 - \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}) + \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}, \end{aligned} \tag{B.29}$$

where (a) comes from Property 3.2 which states that if $Y_{ik} = 1$, then $Y_k = 1$. Then (b) is by

$$\begin{aligned} \mathbb{P}\{Y_k = 1, Y_{ik} = 0 \mid Y_i = 1\} &= \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \cdot \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0, Y_i = 1\} \\ &= \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \cdot \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\} \end{aligned}$$

as Y_i and Y_k are independent knowing that $Y_{ik} = 0$. Similarly,

$$\mathbb{P}\{Y_k = 1\} = \mathbb{P}\{Y_{ik} = 1\} \cdot (1 - \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}) + \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}. \tag{B.30}$$

Moreover,

$$\mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} = \frac{\mathbb{P}\{Y_{ik} = 1, Y_i = 1\}}{\mathbb{P}\{Y_i = 1\}} = \frac{\mathbb{P}\{Y_{ik} = 1\}}{\mathbb{P}\{Y_i = 1\}} \geqslant \mathbb{P}\{Y_{ik} = 1\}. \tag{B.31}$$

From (B.29), (B.30), and (B.31), one obtains that $\mathbb{P}\{Y_k = 1\} \leqslant \mathbb{P}\{Y_k = 1 \mid Y_i = 1\}$. The other inequalities of Lemma B.1 are proved in the same way. □

B.8 Proof of Lemma 3.7

Consider the reference Node i and a node with random index $K \in \mathcal{B}_i \setminus \{i\}$, both known with non-defective sensors. The location of Node K is uniformly distributed on the disk of center i and radius R_0. One has to evaluate

$$\begin{aligned}
&\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\} \\
&= \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(\mathrm{I})}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\}.
\end{aligned}$$

Since $Y_i = 1$, for any realization $\mathbf{m}_{\mathcal{B}_i}$, there exists at least a pair of nodes $(j_1, j_2) \in \mathcal{B}_i^2$ with $j_1 \neq j_2$ such that $[m_{j_1}] \cap [m_{j_2}] = \emptyset$. For the random measurement vector $\mathbf{M}_{\mathcal{B}_i}$, let $\mathcal{D}_i$ the set of such (random) pairs of node indexes. If for some $(J_1, J_2) \in \mathcal{D}_i$, one has also $(J_1, J_2) \in \left(\mathcal{V}_K^{(\mathrm{I})}\right)^2$, then $T\left(\mathbf{M}_{\mathcal{V}_K^{(\mathrm{I})}}\right) = 1$. As a consequence,

$$\begin{aligned}
&\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(\mathrm{I})}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\} \\
&\geqslant \mathbb{P}\left\{J_1 \in \mathcal{V}_K^{(\mathrm{I})}, J_2 \in \mathcal{V}_K^{(\mathrm{I})} \mid (J_1, J_2) \in \mathcal{D}_i, K \in \mathcal{B}_i \setminus \{i\}\right\}.
\end{aligned}$$

One has thus to evaluate the probability that $\mathcal{V}_K^{(\mathrm{I})}$ contains the nodes J_1 and J_2, which themselves belong to $\mathcal{B}_i$. To evaluate this probability, the fact that (J_1, J_2) belong to $\mathcal{D}_i$ is not important, since the outcome of the LODT does not account for the indexes of the nodes. As a consequence

$$\begin{aligned}
&\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(\mathrm{I})}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\} \\
&\geqslant \mathbb{P}\left\{J_1, J_2 \in \mathcal{V}_K^{(\mathrm{I})} \cap \mathcal{B}_i \mid (J_1, J_2) \in \mathcal{D}_i, K \in \mathcal{B}_i \setminus \{i\}\right\} \\
&= \left(\mathbb{P}\left\{J_1 \in \mathcal{V}_K^{(\mathrm{I})} \cap \mathcal{B}_i \mid J_1 \in \mathcal{B}_i, K \in \mathcal{B}_i \setminus \{i\}\right\}\right)^2
\end{aligned} \tag{B.32}$$

since the locations of the nodes are independent.

The distance $R_{i,K}$ between Nodes i and K is a random variable with distribution

$$\pi_{\mathrm{R}}(r) = 2r/R_0^2. \tag{B.33}$$

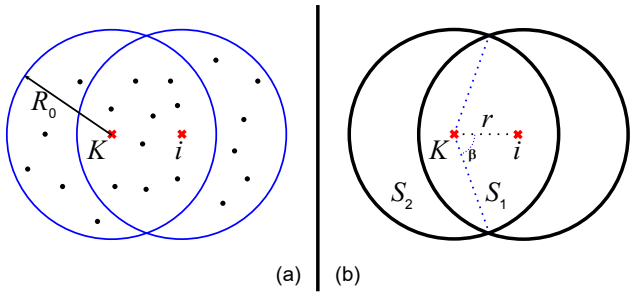

FIGURE B.2: Sensor configuration.

Now, let S_1 be the intersection of the two disks of radius R_0 and centers i and K respectively, see Figure B.2. The probability of a node known to belong to $\mathcal{N}_i$ to be in $\mathcal{U}_i \cap \mathcal{U}_K$ is equal to the area of S_1 divided by πR_0^2, which is a function of $R_{i,K}$. When $R_{i,K} = r$, the area of S_1 is $S_1(\beta) = R_0^2(2\beta - \sin(2\beta))$, with $\beta = \arccos\left(\frac{r}{2R_0}\right)$. Thus, one has to average (B.32) over all possible values of R_{iK} to obtain from (B.33),

$$\begin{aligned}
&\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{N}_i\} \\
&\geqslant \int_0^{R_0} \frac{2r}{R_0^2} \left(\frac{S_1(r)}{\pi R_0^2}\right)^2 \mathrm{d}r \\
&= \int_{\pi/3}^{\pi/2} 4\sin(2\beta) \cdot \left(\frac{2\beta - \sin(2\beta)}{\pi}\right)^2 \mathrm{d}\beta \\
&= 1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2}.
\end{aligned} \tag{B.34}$$

Appendix C

Appendices of Chapter 4

C.1 Proof of Proposition 4.1

For the proof, one considers first the following lemmas.

Lemma C.1. *If*

$$\lim_{t\to\infty}\int_0^t \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau=\infty \tag{C.1}$$

then $p_0p^{00}(t)+p_1p^{10}(t)>0$ *for all* $t\in\mathbb{R}^+$.

Proof. Since $p_0>0$, $p_1>0$, $p^{00}\geqslant 0$, and $p^{10}\geqslant 0$, it suffices to prove that

$$p^{00}(t)+p^{10}(t)\neq 0 \qquad \forall t>0. \tag{C.2}$$

Assume that there exists a time instant $t^*>0$, such that $p^{00}(t^*)+p^{10}(t^*)=0$. As a consequence, at time t^*, all nodes in the network believe themselves as carrying defective sensors. As a consequence, no node will transmit its data to its neighbors. No LODTs will be performed after time t^* and the state of nodes will remain constant. Hence, if $p^{00}(t^*)+p^{10}(t^*)=0$ for some t^*, then $p^{00}(t)+p^{10}(t)=0$ for all $t>t^*$. Consequently,

$$\lim_{t\to\infty}\int_0^t \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau=\int_0^{t^*} \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau,$$

which contradicts (C.1). □

Lemma C.2. *The property* (C.1) *is always satisfied.*

Proof. From (4.32-*a*), one has

$$X_\theta^{0,0}(t)=\exp\left(-\lambda\int_0^t \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau\right). \tag{C.3}$$

Assume that there exists $C^*>0$ such that

$$\lim_{t\to\infty}\int_0^t \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau\leqslant C^* \tag{C.4}$$

then $\forall t\geqslant 0$, one has

$$\int_0^t \left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau\leqslant C^*. \tag{C.5}$$

Combining (C.3) and (C.5), one gets

$$X_\theta^{0,0}(t) \geqslant \exp\left(-\lambda C^*\right) > 0. \tag{C.6}$$

Moreover, from (4.7), one has $p^{00}(\tau) \geqslant X_\theta^{0,0}(\tau)$, leading to

$$\int_0^t \left(p_0 p^{00}(\tau) + p_1 p^{10}(\tau)\right) d\tau \geqslant \int_0^t p_0 X_\theta^{0,0}(\tau) d\tau > p_0 \exp\left(-\lambda C^*\right) t. \tag{C.7}$$

Since $\exp\left(-\lambda C^*\right) t \to \infty$ as $t \to \infty$, (C.7) leads to a violation of the hypothesis (C.4). Hence, one always has (C.1). □

The proof of Proposition 4.1 is then by induction. Starting with (4.32-a), one has (C.3). Since (C.1) is satisfied according to Lemma C.2, for any $\xi > 0$, there exists $t_{00} > 0$ such that $t > t_{00}$ implies $X_\theta^{0,0}(t) < \xi$ and $\lim_{t\to\infty} X_\theta^{0,0}(t) = 0$.

Then, assume that for any $\ell \leqslant M-1$, and $\xi > 0$, there exists $t_{(\ell-1)0} > \cdots > t_{00}$ such that $t > t_{(\ell-1)0}$ implies $X_\theta^{j,0}(t) < \xi$ for $j = 0, \ldots, \ell-1$. One has to show now that there exists $t_{\ell 0} > t_{(\ell-1)0}$ such that $X_\theta^{\ell,0}(t) < \xi$ for all $t > t_{\ell 0}$.

Define $Z_\theta^{\ell,0}(t) = \sum_{j=0}^{\ell} X_\theta^{j,0}(t)$. From (4.32$a$) and (4.32$b$), one has

$$\frac{dZ_\theta^{\ell,0}}{dt} = -\lambda \left(v(t) Z_\theta^{\ell-1,0}(t) + \left(p_0 p^{00}(t) + p_1 p^{10}(t)\right) X_\theta^{\ell,0}(t) \right),$$

where $v(t) = \pi_\theta^{1,1}(t, \ell, k)$, since $\pi_\theta^{1,0}$ and $\pi_\theta^{1,1}$ do not depend on ℓ and k when $\ell < M$. Using (C.2) one has $dZ_\theta^{\ell,0}/dt < 0$ for any $X_\theta^{\ell 0} > 0$. As a consequence, $Z_\theta^{\ell,0}(t)$ decreases until $X_\theta^{\ell,0}(t)$ reaches 0. Hence, for any $\xi > 0$, there exists $t_{\ell,0} > t_{(\ell-1)0}$, such that $X_\theta^{\ell,0} < \xi$ and $\lim_{t\to\infty} X_\theta^{\ell,0}(t) = 0$.

In the same way, using (4.32c) and the previous results that $X_\theta^{\ell k}(t) \to 0$ with $k = 1, \ldots, M-2$ and $\ell = k, \ldots, M-2$, one can prove that for any $k = 1, \ldots, M-1$, $X_\theta^{\ell',(k+1)}(t)$ tends to zero as $t \to \infty$, with any $\ell' = k+1, \ldots, M-1$.

C.2 Proof of Proposition 4.2

According to Proposition 4.1, one has $\overline{X}_\theta^{\ell,k} = 0$, for all $\ell < M$ and $k \leqslant \ell$. To evaluate $\overline{X}_\theta^{M,k}$, one thus considers the following simplified dynamics derived from (4.32) for $\theta \in \{0, 1\}$,

$$\begin{cases} \frac{dX_\theta^{M,0}}{dt} = \lambda\left(-X_\theta^{M,0}\pi_\theta^{0,1}(M,0) + X_\theta^{M,1}\pi_\theta^{0,-1}(M,1)\right), \\ \frac{dX_\theta^{M,M}}{dt} = \lambda\left(-X_\theta^{M,M}\pi_\theta^{0,-1}(M,M) + X_\theta^{M,M-1}\pi_\theta^{0,1}(M,M-1)\right), \\ \frac{dX_\theta^{M,k}}{dt} = \lambda\left(-X_\theta^{M,k}\left(\pi_\theta^{0,-1}(M,k) + \pi_\theta^{0,1}(M,k)\right) + X_\theta^{M,k+1}\pi_\theta^{0,-1}(M,k+1) + X_\theta^{M,k-1}\pi_\theta^{0,1}(M,k-1)\right) \end{cases} \tag{C.8}$$

At equilibrium, one has $dX_\theta^{M,k}(t)/dt = 0$ for all $k \leqslant M$. Moreover, the transition probabilities do not vary any more. Let $\overline{\boldsymbol{X}}_\theta^M = \left(\overline{X}_\theta^{M,1}, \ldots, \overline{X}_\theta^{M,M}\right)^{\mathrm{T}}$, $a_\theta(k) = \pi_\theta^{0,1}(M,k)$, and $b_\theta(k) = \pi_\theta^{0,-1}(M,k)$.

From (C.8), one deduces that the vector $\overline{\boldsymbol{X}}_\theta^M$ should satisfy $\boldsymbol{\Psi}_\theta \cdot \overline{\boldsymbol{X}}_\theta^M = \mathbf{0}$ where

$$\boldsymbol{\Psi}_\theta = \begin{pmatrix} -a_\theta(0) & b_\theta(1) & & \\ a_\theta(0) & -a_\theta(1) - b_\theta(1) & b_\theta(2) & \\ & \ddots & \ddots & \ddots \\ & & a_\theta(M-1) & -b_\theta(M) \end{pmatrix}.$$

Summing Lines 1 to $k+1$, for all $k = 0, \ldots, M-1$, one obtains $a_\theta(k)\,\overline{X}_\theta^{M,k} = b_\theta(k+1)\,\overline{X}_\theta^{M,k+1}$, which leads to

$$\overline{X}_\theta^{M,k} = \overline{X}_\theta^{M,0} \prod_{j=0}^{k-1} \frac{a_0(j)}{b_0(j+1)}. \tag{C.9}$$

One evaluates

$$\frac{a_\theta(j)}{b_\theta(j+1)} = \frac{\pi_\theta^{0,1}(M,j)}{\pi_\theta^{0,-1}(M,j+1)} = \eta_\theta \frac{M-j}{j+1}, \tag{C.10}$$

where using (4.26) and (4.27), one has

$$\begin{cases} \eta_0 = \frac{p_0 q_{\mathrm{FA}}(2)\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\overline{p}^{10}}{p_0(1-q_{\mathrm{FA}}(2))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(1,1))\overline{p}^{10}}, \\ \eta_1 = \frac{p_0 q_{\mathrm{D}}(1,1)\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\overline{p}^{10}}{p_0(1-q_{\mathrm{D}}(1,1))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(0,2))\overline{p}^{10}}. \end{cases} \tag{C.11}$$

with $\overline{p}^{00}$ and $\overline{p}^{10}$ defined in (4.33).

From (C.9) and (C.10), one deduces

$$\begin{aligned} \overline{X}_\theta^{M,k} &= \overline{X}_\theta^{M,0} \prod_{j=0}^{k-1} \left(\eta_\theta \frac{M-j}{j+1} \right) \\ &= \overline{X}_\theta^{M,0} \eta_\theta^k \frac{M \cdot (M-1) \cdot (M-k+1)}{1 \cdot 2 \cdot \cdots \cdot k} = \binom{M}{k} \eta_\theta^k \overline{X}_\theta^{M,0}. \end{aligned} \tag{C.12}$$

Since $\sum_{k=0}^M \overline{X}_\theta^{M,k} = 1$, one has

$$1 = \sum_{k=0}^{M} \binom{M}{k} \eta_\theta^k \overline{X}_\theta^{M,0} = (\eta_\theta + 1)^M \overline{X}_\theta^{M,0}. \tag{C.13}$$

From (C.12) and (C.13), $\forall k = 0, \ldots, M$,

$$\overline{X}_\theta^{M,k} = \binom{M}{k} \left(\frac{\eta_\theta}{\eta_\theta + 1} \right)^k \left(\frac{1}{\eta_\theta + 1} \right)^{M-k} = \binom{M}{k} (h_\theta)^k (1-h_\theta)^{M-k} \tag{C.14}$$

with $h_\theta = \frac{\eta_\theta}{\eta_\theta + 1}$. Introducing (C.14) into (4.33), one obtains (4.38) with F_θ defined in (4.37). Thus one needs to solve (4.38) to determine $\overline{\mathbf{p}}$, which is then used to deduce $\overline{X}_\theta^{M,d}$ using (C.14).

C.3 Proof of Lemma 4.1

To prove Lemma 4.1, one needs first investigate the monotonicity of F_θ. To lighten the notations, let $\alpha = q_{\mathrm{FA}}(2)$, $\beta = q_{\mathrm{D}}(1,1)$ and $\gamma = q_{\mathrm{D}}(0,2)$. Then h_0 and h_1 defined in (4.35-4.36) can be rewritten as

$$h_0(x,y) = \frac{\alpha p_0 x + \beta p_1 y}{p_0 x + p_1 y},\ h_1(x,y) = \frac{\beta p_0 x + \gamma p_1 y}{p_0 x + p_1 y}, \tag{C.15}$$

with $(x,y) \in \mathcal{P}_0$. One starts showing some monotonicity properties.

Lemma C.3. *If $\alpha < \beta < \gamma$, then h_0 and h_1 are decreasing with x and increasing with y, for all $(x,y) \in \mathcal{P}_0$. If $\beta = \gamma$, then $h_1 = \beta = \gamma$ is a constant.*

Proof. Since $\alpha < \beta \leqslant \gamma$, one has

$$\begin{aligned}
\frac{\partial h_0}{\partial x} &= \frac{(\alpha-\beta)\, p_0 p_1 y}{(p_0 x + p_1 y)^2} \leqslant 0, \quad \frac{\partial h_0}{\partial y} = \frac{(\beta-\alpha)\, p_0 p_1 x}{(p_0 x + p_1 y)^2} \geqslant 0, \\
\frac{\partial h_1}{\partial x} &= \frac{(\beta-\gamma)\, p_0 p_1 y}{(p_0 x + p_1 y)^2} \leqslant 0, \quad \frac{\partial h_1}{\partial y} = \frac{(\gamma-\beta)\, p_0 p_1 x}{(p_0 x + p_1 y)^2} \geqslant 0.
\end{aligned}$$

then Lemma C.3 can be proved. □

Lemma C.4. *For $z \in [0,1]$, the family of functions*

$$f_i(z) = z^i (1-z)^{M-i},\ i = 0, 1 \ldots, M. \tag{C.16}$$

are increasing over $[0, \frac{i}{M}]$ and decreasing over $[\frac{i}{M}, 1]$.

Proof. Consider three possible situations: 1) If $i = 0$, $f_0(z) = (1-z)^M$ is decreasing over $[0,1]$. 2) If $i = M$, $f_M(z) = z^M$ is increasing over $[0,1]$. 3) If $1 \leqslant i \leqslant M-1$,

$$\frac{df_i}{dz} = z^{i-1}(1-z)^{M-i-1}(i - Mz), \tag{C.17}$$

and $df_i/dz \geqslant 0$ when $z \in [0, \frac{i}{M}]$ and $df_i/dz \leqslant 0$ when $z \in [\frac{i}{M}, 1]$. Therefore, Lemma C.4 holds $\forall i = 0, \ldots, M$. □

Lemma C.5. *If $0 < \nu < 1$, the function*

$$g(z) = \sum_{i:i/M<\nu} \binom{M}{i} f_i(z) = \sum_{i:i/M<\nu} \binom{M}{i} z^i (1-z)^{M-i}, \tag{C.18}$$

is decreasing for all $z \in [0,1]$.

Proof. First, consider $z \in [\nu, 1]$. In (C.18), each i in the sum is such that $\frac{i}{M} < \nu \leqslant z$. From Lemma C.4, $f_i(z)$ is a decreasing function for any $\frac{i}{M} < z$, thus $g(z)$ is also decreasing with z.

Now, consider $z \in [0, \nu]$, one rewrites (C.18) as

$$g(z) = 1 - \sum_{i:i/M\geqslant\nu} \binom{M}{i} f_i(z), \tag{C.19}$$

in which each i in the sum is such that $z < \nu \leqslant \frac{i}{M}$. Applying again Lemma C.4, since $f_i(z)$ is an increasing function for any $z \leqslant \frac{i}{M}$, the sum in (C.19) is also increasing with z and $g(z)$ is decreasing. Thus $g(z)$ is a decreasing function of z over $[0,1]$. □

Considering the functions h_θ and g, then one may rewrite F_θ as $F_\theta(x,y) = g(h_\theta(x,y))$, $\forall\theta \in \{0,1\}$. The monotonicity of F_0 and F_1 is shown in the following lemma.

Lemma C.6. *If $\alpha < \beta < \gamma$, then F_0 and F_1 are increasing functions of x and decreasing functions of y, for all $(x,y) \in \mathcal{P}_0$. If $\beta = \gamma$, then $F_1 = g(\beta) = g(\gamma)$ is a constant.*

Proof. The proof of obtained by combining Lemma C.3 and Lemma C.5. □

The proof of Lemma 4.1 is by induction. At the beginning, one has $0 \leqslant p^{\theta 0}(0) \leqslant 1$, thus $p^{\theta 0}_{\min}(0) = 0$ and $p^{\theta 0}_{\max}(0) = 1$. Using Lemma C.6, one has $F_\theta(0,1) \leqslant F_\theta\left(p^{00}(0), p^{10}(0)\right) \leqslant F_\theta(1,0)$, thus

$$\begin{cases} p^{00}_{\min}(1) = F_0(0,1) = g(\beta) > 0 = p^{00}_{\min}(0), \\ p^{00}_{\max}(1) = F_0(1,0) = g(\alpha) < 1 = p^{00}_{\max}(0), \\ p^{10}_{\min}(1) = F_1(0,1) = g(\gamma) > 0 = p^{10}_{\min}(0), \\ p^{10}_{\max}(1) = F_1(1,0) = g(\beta) < 1 = p^{10}_{\max}(0), \end{cases} \tag{C.20}$$

thus (4.42) and (4.43) are true for $n = 1$.

Consider then an arbitrary $n \in \mathbb{N}^*$ and $n > 1$. Assume that (4.42) and (4.43) are satisfied for any $n' < n$ and $n' \in \mathbb{N}^*$, one needs to see whether (4.42) and (4.43) are still satisfied for n. Applying Lemma C.6 again, one obtains

$$\begin{aligned} p^{\theta 0}_{\min}(n) &= F_\theta\left(p^{00}_{\min}(n-1), p^{10}_{\max}(n-1)\right) \\ &> F_\theta\left(p^{00}_{\min}(n-2), p^{10}_{\max}(n-2)\right) = p^{\theta 0}_{\min}(n-1), \end{aligned}$$

and

$$\begin{aligned} p^{\theta 0}_{\max}(n) &= F_\theta\left(p^{00}_{\max}(n-1), p^{10}_{\min}(n-1)\right) \\ &< F_\theta\left(p^{00}_{\max}(n-2), p^{10}_{\min}(n-2)\right) = p^{\theta 0}_{\max}(n-1), \end{aligned}$$

Similarly, one gets $p^{10}_{\min}(n) > p^{10}_{\min}(n-1)$ and $p^{\theta 0}_{\max}(n) < p^{\theta 0}_{\max}(n-1)$.

C.4 Proof of Proposition 4.4

As seen in the proof of Proposition 4.3, $\forall n \in \mathbb{N}^*$, $\mathbf{F}(\mathbf{p})$ maps $\mathcal{P}_n$ to $\mathcal{P}_n$, with $\mathcal{P}_n = \left[p^{00}_{\min}(n), p^{00}_{\max}(n)\right] \times \left[p^{10}_{\min}(n), p^{10}_{\max}(n)\right]$. In order to apply Banach's fixed-point theorem [Ban22] to prove Proposition 4.4, it suffices to show that $\mathbf{F}$ is contracting, *i.e.*, that for any pairs $\mathbf{p} = (x,y) \in \mathcal{P}_n$ and $\mathbf{p} + \boldsymbol{\delta} = (x + \delta_x, y + \delta_y) \in \mathcal{P}_n$, one has

$$|\mathbf{F}(\mathbf{p} + \boldsymbol{\delta}) - \mathbf{F}(\mathbf{p})| < |\boldsymbol{\delta}|. \tag{C.21}$$

A sufficient condition to have (C.21) is that the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} \frac{\partial F_0(x,y)}{\partial x} & \frac{\partial F_0(x,y)}{\partial y} \\ \frac{\partial F_1(x,y)}{\partial x} & \frac{\partial F_1(x,y)}{\partial y} \end{pmatrix}$$

have module less than 1. The eigenvalues of $\mathbf{A}$ are the solutions of

$$z^2 - \left(\frac{\partial F_0}{\partial x} + \frac{\partial F_1}{\partial y}\right) z + \left(\frac{\partial F_0}{\partial x}\frac{\partial F_1}{\partial y} - \frac{\partial F_0}{\partial y}\frac{\partial F_1}{\partial x}\right) = 0. \tag{C.22}$$

As in Appendix C.3, denote $\alpha = q_{\mathrm{FA}}(2)$, $\beta = q_{\mathrm{D}}(1,1)$ and $\gamma = q_{\mathrm{D}}(0,2)$. First, one evaluates

$$\frac{\partial F_0}{\partial x}\frac{\partial F_1}{\partial y} - \frac{\partial F_0}{\partial y}\frac{\partial F_1}{\partial x} = \frac{\partial g}{\partial h_0}\frac{\partial g}{\partial h_1}\left(\frac{\partial h_0}{\partial x}\frac{\partial h_1}{\partial y} - \frac{\partial h_0}{\partial y}\frac{\partial h_1}{\partial x}\right) \overset{(a)}{=} 0,$$

where (a) comes from $\frac{\partial h_0}{\partial x}\frac{\partial h_1}{\partial y} = \frac{\partial h_0}{\partial y}\frac{\partial h_1}{\partial x}$, using the partial derivatives calculated in the proof of Lemma C.3. Then, the solutions of (C.22) are $z_1 = \frac{\partial F_0}{\partial x} + \frac{\partial F_1}{\partial y}$ and $z_2 = 0$. Hence, it suffices to prove that $|z_1| < 1$.

We begin with the evaluation of an upper bound of the partial derivative of $F_0(x,y)$ with respect to x

$$\begin{aligned} \frac{\partial F_0(x,y)}{\partial x} &= \frac{\partial g(h_0(x,y))}{\partial x} = \frac{\partial g}{\partial h_0}\cdot\frac{\partial h_0}{\partial x} \\ &\overset{(a)}{=} \frac{(\beta-\alpha)\,p_0p_1y}{(p_0x+p_1y)^2}\sum_{i:i/M<\nu}\binom{M}{i}h_0^i(1-h_0)^{M-i}\frac{h_0M-i}{h_0(1-h_0)} \\ &\overset{(b)}{\leqslant} \frac{(\beta-\alpha)\,p_0p_1y}{(p_0x+p_1y)^2}F_0(x,y)\frac{M}{1-h_0} \leqslant c_0(\alpha,\beta,\gamma,M,\nu,n), \end{aligned} \tag{C.23}$$

where (a) is obtained using (C.17), (b) comes from $i \geqslant 0$, and c_0 is defined in (4.45). Meanwhile, from Lemma C.6, one has $\partial F_0(x,y)/\partial x \geqslant 0$, as F_0 is an increasing function of x.

Similarly,

$$\begin{aligned} \frac{\partial F_1(x,y)}{\partial x} &= \frac{\partial g(h_1(x,y))}{\partial y} = \frac{\partial g}{\partial h_1}\cdot\frac{\partial h_1}{\partial y} \\ &= \frac{(\gamma-\beta)\,p_0p_1x}{(p_0x+p_1y)^2}\sum_{i:i/M<\nu}\binom{M}{i}h_1^i(1-h_1)^{M-i}\frac{i-h_1M}{h_1(1-h_1)} \\ &\geqslant \frac{(\gamma-\beta)\,p_0p_1x}{(p_0x+p_1y)^2}F_1(x,y)\frac{-M}{1-h_1} \geqslant -c_1(\alpha,\beta,\gamma,M,\nu,n), \end{aligned} \tag{C.24}$$

and $\partial F_1(x,y)/\partial y \leqslant 0$ as F_1 is a non-decreasing function of y. One concludes that

$$-c_1 \leqslant \frac{\partial F_0(x,y)}{\partial x} + \frac{\partial F_1(x,y)}{\partial y} \leqslant c_0,$$

thus $c_0 < 1$ and $c_1 < 1$ lead to $|z_1| < 1$, which ensures the unicity of the equilibrium.

C.5 Proof of Proposition 4.5

First, one shows that if $\nu < q_{\mathrm{D}}(1,1)$, then for any $\varepsilon > 0$, there exists $M > M'$, such that $\overline{p}^{10} < \varepsilon$.

From Lemma 4.1, $\overline{p}^{10}$ can be bounded as

$$\overline{p}^{10} = F_1\left(\overline{p}^{00}, \overline{p}^{10}\right) < \sum_{k:k/M<\nu} \binom{M}{k} \left(q_{\mathrm{D}}(1,1)\right)^k \left(1 - q_{\mathrm{D}}(1,1)\right)^{M-k} \tag{C.25}$$

Consider $\Phi_1, \Phi_2, \ldots$ an infinite sequence of *i.i.d.* binary random variables with $\mathbb{P}\{\Phi_m = 1\} = q_{\mathrm{D}}(1,1)$. For any $\varrho \in [0,1]$ such that $\varrho M \in \mathbb{N}^+$, one has

$$\mathbb{P}\left\{\frac{\sum_{m=1}^{M} \Phi_m}{M} = \varrho\right\} = \binom{M}{\varrho M} \left(q_{\mathrm{D}}(1,1)\right)^{\varrho M} \left(1 - q_{\mathrm{D}}(1,1)\right)^{M(1-\varrho)}.$$

According to the weak law of large numbers [CT12], for $\varepsilon > 0$, there exists M', such that for any $M > M'$, one has

$$\mathbb{P}\left\{\left|\frac{\sum_{m=1}^{M} \Phi_m}{M} - q_{\mathrm{D}}(1,1)\right| > q_{\mathrm{D}}(1,1)\right\} < \varepsilon. \tag{C.26}$$

From (C.26), one also has

$$\begin{aligned}\sum_{k:k/M<(q_{\mathrm{D}}(1,1)-\varepsilon)} \left(q_{\mathrm{D}}(1,1)\right)^k \left(1 - q_{\mathrm{D}}(1,1)\right)^{M-k} &= \mathbb{P}\left\{\frac{\sum_{m=1}^{M} \Phi_m}{M} - q_{\mathrm{D}}(1,1) < -\varepsilon\right\} \\ &\leqslant \mathbb{P}\left\{\left|\frac{\sum_{m=1}^{M} \Phi_m}{M} - q_{\mathrm{D}}(1,1)\right| > \varepsilon\right\} < \varepsilon. \end{aligned} \tag{C.27}$$

If $\nu < q_{\mathrm{D}}(1,1) - \varepsilon$, then using (C.27), the bound of $\overline{p}^{10}$ in (C.25) may be further written as

$$\begin{aligned}\overline{p}^{10} &< \sum_{k:k/M<\nu} \binom{M}{k} \left(q_{\mathrm{D}}(1,1)\right)^k \left(1 - q_{\mathrm{D}}(1,1)\right)^{M-k} \\ &\leqslant \sum_{k:k/M<(q_{\mathrm{D}}(1,1)-\varepsilon)} \binom{M}{k} \left(q_{\mathrm{D}}(1,1)\right)^k \left(1 - q_{\mathrm{D}}(1,1)\right)^{M-k} < \varepsilon. \end{aligned} \tag{C.28}$$

From Lemma C.3 and the fact that $q_{\mathrm{FA}}(2) \leqslant \overline{p}^{10} \leqslant q_{\mathrm{D}}(1,1)$ and $0 \leqslant \overline{p}^{10} < \varepsilon$, one has $h_0\left(\overline{p}^{00}, \overline{p}^{10}\right) \in [q_{\mathrm{FA}}(2), \chi(\varepsilon)]$, with

$$\chi(\varepsilon) = \frac{p_0 \left(q_{\mathrm{FA}}(2)\right)^2 + p_1 q_{\mathrm{D}}(1,1)\,\varepsilon}{p_0 q_{\mathrm{FA}}(2) + p_1 \varepsilon}. \tag{C.29}$$

Thus, according to Lemma C.5,

$$\begin{aligned}\overline{p}^{00} &= F_0\left(\overline{p}^{00}, \overline{p}^{10}\right) = g\left(h_0\left(\overline{p}^{00}, \overline{p}^{10}\right)\right) \\ &\geqslant g(\chi(\varepsilon)) = \sum_{k:k/M<\nu} \binom{M}{k} \left(\chi(\varepsilon)\right)^k \left(1 - \chi(\varepsilon)\right)^{M-k}. \end{aligned} \tag{C.30}$$

Using derivations similar to those leading to (C.27), one gets

$$\sum_{k:k/M>(\chi(\varepsilon)+\varepsilon)} \binom{M}{k} (\chi(\varepsilon))^k (1-\chi(\varepsilon))^{M-k} < \varepsilon, \tag{C.31}$$

which leads to

$$\sum_{k:k/M\leqslant(\chi(\varepsilon)+\varepsilon)} \binom{M}{k} (\chi(\varepsilon))^k (1-\chi(\varepsilon))^{M-k} \geqslant 1-\varepsilon. \tag{C.32}$$

If $\nu > \chi(\varepsilon) + \varepsilon$, then

$$\begin{aligned} \overline{p}^{00} &\geqslant \sum_{k:k/M<\nu} \binom{M}{k} (\chi(\varepsilon))^k (1-\chi(\varepsilon))^{M-k} \\ &\geqslant \sum_{k:k/M\leqslant(\chi(\varepsilon)+\varepsilon)} \binom{M}{k} (\chi(\varepsilon))^k (1-\chi(\varepsilon))^{M-k} \geqslant 1-\varepsilon. \end{aligned} \tag{C.33}$$

As a conclusion, for any $\varepsilon > 0$, if $\chi(\varepsilon) + \varepsilon < \nu < q_{\mathrm{D}}(1,1) - \varepsilon$, then $\overline{p}^{00} \geqslant 1-\varepsilon$ and $\overline{p}^{10} < \varepsilon$. Since $\lim_{\varepsilon\to 0} \chi(\varepsilon) = q_{\mathrm{FA}}(2)$, one concludes that if $q_{\mathrm{FA}}(2) < \nu < q_{\mathrm{D}}(1,1)$, one obtains (4.47).

Appendix D

Appendices of Chapter 5

D.1 Proof of Proposition 5.2

Consider an arbitrary $\theta \in \Theta$ and some $\varepsilon > 0$ such that

$$\varepsilon < \min_{1 \leqslant \theta < K} \frac{s_\theta - s_{\theta+1}}{2}. \tag{D.1}$$

Let $\Phi_1^\theta, \Phi_2^\theta, \ldots$ be an infinite sequence of *i.i.d.* binary random variables with $\mathbb{P}\left\{\Phi_m^\theta = 1\right\} = s_\theta$. For any $\varrho \in [0,1]$ such that $\varrho M \in \mathbb{N}^+$, one has

$$\mathbb{P}\left\{\frac{\sum_{m=1}^M \Phi_m^\theta}{M} = \varrho\right\} = \binom{M}{\varrho M} (s_\theta)^{\varrho M} (1 - s_\theta)^{M(1-\varrho)}. \tag{D.2}$$

According to the weak law of large numbers [CT12], there exists M', such that for all $M > M'$, one has

$$\mathbb{P}\left\{\left|\frac{\sum_{m=1}^M \Phi_m^\theta}{M} - s_\theta\right| < \varepsilon\right\} > 1 - \varepsilon. \tag{D.3}$$

From (D.2) and (D.3), one also has

$$\begin{aligned}
&\sum_{\beta:\beta/M \in (s_\theta - \varepsilon, s_\theta + \varepsilon)} \binom{M}{\beta} (s_\theta)^\beta (1 - s_\theta)^{M-\beta} \\
&= \mathbb{P}\left\{\frac{\sum_{m=1}^M \Phi_m^\theta}{M} \in (s_\theta - \varepsilon, s_\theta + \varepsilon)\right\} \\
&> 1 - \varepsilon.
\end{aligned} \tag{D.4}$$

If $\nu_{\theta-1}$ and ν_θ are chosen such that

$$\nu_{\theta-1} > s_\theta + \varepsilon \tag{D.5}$$

and

$$s_\theta - \varepsilon > \nu_\theta, \tag{D.6}$$

then $\forall \theta \in \Theta$,

$$\begin{aligned}\overline{p}^{\theta\theta} &= \sum_{\beta:\beta/M\in(\nu_\theta,\nu_{\theta-1})} \binom{M}{\beta} (s_\theta)^\beta (1-s_\theta)^{M-\beta} \\ &> \sum_{\beta:\beta/M\in(s_\theta-\varepsilon,s_\theta+\varepsilon)} \binom{M}{\beta} (s_\theta)^\beta (1-s_\theta)^{M-\beta} > 1-\varepsilon.\end{aligned}$$

The constraints (D.5) and (D.6) have to be satisfied for $\theta = 1,\dots,K$. They may be reformulated as

$$s_{\theta+1} + \varepsilon < \nu_\theta < s_\theta - \varepsilon \tag{D.7}$$

for $\theta = 1,\dots,K-1$. Such values of ν_θ may be found, since one imposes the constraints (D.1) on ε.

One concludes that $\forall \theta \in \Theta$ and $\forall \varepsilon > 0$ satisfying (D.1), there exists a choice of the values of ν_θ such that (D.7) is satisfied and there exists M', such that for all $M > M'$, $\overline{p}^{\theta\theta} > 1-\varepsilon$. The value of ε can be chosen arbitrarily close to zero, which will require M going to infinity, so, provided that $s_{\theta+1} < \nu_\theta < s_\theta$, $\forall \theta \in \Theta$, one has $\lim_{M\to\infty} \overline{p}^{\theta\theta} = 1$, if $\nu_\theta < s_\theta < \nu_{\theta-1}$, $\forall \theta \in \Theta$.

D.2 Proof of Proposition 5.4

Brouwer's fixed-point theorem [GD13] can be used to show the existence of a solution of (5.24). For that purpose, one has to show that for any $\mathbf{p}^1(0) = \left(p^{11}(0)\dots p^{K1}(0)\right) \in \mathcal{P}_0 = \left\{\boldsymbol{x} \in [0,1]^K \text{ and } \boldsymbol{x} \neq \mathbf{0}\right\}$, the discrete-time system

$$\mathbf{p}^1(n+1) = \mathbf{F}\left(\mathbf{p}^1(n)\right), \tag{D.8}$$

with $\mathbf{p}^1(n) = \left(p^{11}(n)\dots p^{K1}(n)\right)$ $\forall n \in \mathbb{N}^*$, converges to a equilibrium point $\overline{\mathbf{p}}^1$.

Each F_θ with $\theta \in \Theta$ is a continuous function. One needs to verify whether the value $\mathbf{p}^1(0)$ belongs to some compact set. Using the similar derivations as in the proofs of Lemma C.1 and C.2, one obtains that there is always

$$\sum_{\theta\in\Theta} p^{\theta 1}(t) > 0. \tag{D.9}$$

which means that

$$\mathbf{p}^1(0) \in \mathcal{P}_0 = \left\{\boldsymbol{x} \in [0,1]^K \text{ such that } \boldsymbol{x} \neq \mathbf{0}\right\}. \tag{D.10}$$

$\mathcal{P}_0$ is not compact. One then has to find a compact $\mathcal{P}_n$ such that $\mathbf{F}$ maps $\mathcal{P}_n$ to $\mathcal{P}_n$, in order to apply Brouwer's fixed-point. One starts showing some properties of some basic functions involved in $\mathbf{F}$.

Lemma D.1. *If $\boldsymbol{x} = (x_1\dots x_K) \in \mathcal{P}_0$, then $h_\theta(\boldsymbol{x})$ is bounded as follows*

$$0 < h_{\theta,\min} \leqslant h_\theta(\boldsymbol{x}) \leqslant h_{\theta,\max}, \tag{D.11}$$

where

$$\begin{cases} h_{\theta,\max} = \max\left\{q\left(\theta,1\right)\ldots q\left(\theta,K\right)\right\}, \\ h_{\theta,\min} = \min\left\{q\left(\theta,1\right)\ldots q\left(\theta,K\right)\right\}. \end{cases} \tag{D.12}$$

Proof. Using Assumption A6), one has $h_{\theta,\min} = \min\left\{q\left(\theta,1\right)\ldots q\left(\theta,K\right)\right\} > 0$. Moreover, one has

$$\begin{aligned} h_{\theta,\max} - h_\theta\left(\boldsymbol{x}\right) &= h_{\theta,\max} - \frac{\sum_{k\in\Theta} p_k q\left(\theta,k\right) x_k}{\sum_{k\in\Theta} p_k x_k} \\ &= \frac{\sum_{k\in\Theta} p_k \left(h_{\theta,\max} - q\left(\theta,k\right)\right) x_k}{\sum_{k\in\Theta} p_k x_k}. \end{aligned}$$

Since $h_{\theta,\max} \geqslant q\left(\theta,k\right)$, $\forall k \in \Theta$, one gets $h_{\theta,\max} - h_\theta\left(\boldsymbol{x}\right) \geqslant 0$. In a similar way,

$$h_{\theta,\min} - h_\theta\left(\boldsymbol{x}\right) = \frac{\sum_{k\in\Theta} p_k \left(h_{\theta,\min} - q\left(\theta,k\right)\right) x_k}{\sum_{k\in\Theta} p_k x_k} \leqslant 0.$$

Then (D.11) is proved. □

Lemma D.2. *If* $0 < \nu_1 < 1$*, the function*

$$g(z) = \sum_{\beta=\lceil M\nu_1\rceil}^{M} \binom{M}{\beta} z^\beta \left(1-z\right)^{M-\beta}, \tag{D.13}$$

is increasing for all $z \in [0,1]$.

Proof. In Lemma C.5, one has shown that $1-g(z)$ is decreasing for $z \in [0,1]$, thus $g(z)$ is increasing. □

From Lemma D.1 and Lemma D.2, one obtains that for any $\theta \in \Theta$ and $\boldsymbol{x} \in \mathcal{P}_0$

$$\begin{cases} F_\theta\left(\boldsymbol{x}\right) \leqslant g\left(h_{\theta,\max}\right), \\ F_\theta\left(\boldsymbol{x}\right) \geqslant g\left(h_{\theta,\min}\right) > 0 \end{cases} \tag{D.14}$$

Define $p^{\theta 1}_{\max}\left(n\right)$ and $p^{\theta 1}_{\min}\left(n\right)$ as upper and lower bounds of $p^{\theta 1}\left(n\right)$, *i.e.*, $p^{\theta 1}_{\min}\left(n\right) \leqslant p^{\theta 1}\left(n\right) \leqslant p^{\theta 1}_{\max}\left(n\right)$. When $n = 0$, one has $p^{\theta 1}_{\min}\left(0\right) = 0$ and $p^{\theta 0}_{\max}\left(0\right) = 1$. From (D.14), one gets

$$\begin{cases} p^{\theta 1}_{\min}\left(1\right) = g\left(h_{\theta,\min}\right) > 0 \\ p^{\theta 1}_{\max}\left(1\right) = g\left(h_{\theta,\max}\right). \end{cases} \tag{D.15}$$

Define

$$\mathcal{P}_1 = \left[p^{11}_{\min}\left(1\right), p^{11}_{\max}\left(1\right)\right] \times \ldots \times \left[p^{K1}_{\min}\left(1\right), p^{K1}_{\max}\left(1\right)\right], \tag{D.16}$$

then $\mathbf{p}^1\left(1\right) = \mathbf{F}\left(\mathbf{p}^1\left(0\right)\right) \in \mathcal{P}_1$. Notice that $\mathbf{0} \notin \mathcal{P}_1$ and $\mathcal{P}_1$ is a compact set since $p^{\theta 1}_{\min}\left(1\right) > 0$

Consider then an arbitrary integer $n \in \mathbb{N}^*$. Assume that $\mathbf{p}^1\left(n-1\right) \in \mathcal{P}_1$, one needs to see whether $\mathbf{p}^1\left(n\right) \in \mathcal{P}_1$ is satisfied. Since $\mathcal{P}_1 \subseteq \mathcal{P}_0$, one still has

$$h_{\theta,\min} \leqslant h_\theta\left(\mathbf{p}^1\left(n-1\right)\right) \leqslant h_{\theta,\max},$$

which leads to

$$\begin{cases} p^{\theta 1}(n) = F_\theta\left(\mathbf{p}^1(n-1)\right) \leqslant g\left(h_{\theta,\max}\right) \\ p^{\theta 1}(n) = F_\theta\left(\mathbf{p}^1(n-1)\right) \geqslant g\left(h_{\theta,\min}\right) \end{cases}. \tag{D.17}$$

Therefore $\mathbf{F}$ maps $\mathcal{P}_1$ to $\mathcal{P}_1$. Besides, $\mathcal{P}_1$ is compact. Hence one can apply Brouwer's fixed-point theorem to prove Proposition 5.4.

D.3 Proof of Proposition 5.5

First, one investigates the upper bound of $\overline{p}^{\theta 1}$ for $\theta \in \Theta \setminus \{1\}$ and M sufficient large. As for (D.4), for any $\varepsilon > 0$, there exists m_1, such that for all $M > m_1$

$$\sum_{\beta=\lfloor (s_\theta-\varepsilon)M \rfloor+1}^{\lceil (s_\theta+\varepsilon)M \rceil-1} \binom{M}{\beta} \left(h_{\theta,\max}\right)^\beta \left(1-h_{\theta,\max}\right)^{M-\beta} > 1-\varepsilon, \tag{D.18}$$

which leads to

$$\begin{aligned} &\sum_{\beta=\lfloor (h_{\theta,\max}+\varepsilon)M \rfloor+1}^{M} \binom{M}{\beta} \left(h_{\theta,\max}\right)^\beta \left(1-h_{\theta,\max}\right)^{M-\beta} \\ &< 1 - \sum_{\beta=\lfloor (s_\theta-\varepsilon)M \rfloor+1}^{\lceil (s_\theta+\varepsilon)M \rceil-1} \binom{M}{\beta} \left(h_{\theta,\max}\right)^\beta \left(1-h_{\theta,\max}\right)^{M-\beta} \\ &< \varepsilon. \end{aligned} \tag{D.19}$$

If $\nu_1 > h_{\theta,\max} + \varepsilon$, then from (D.17) and (D.19), one gets

$$\begin{aligned} \overline{p}^{\theta 1} &< \sum_{\beta=\lceil \nu_1 M \rceil}^{M} \binom{M}{\beta} \left(h_{\theta,\max}\right)^\beta \left(1-h_{\theta,\max}\right)^{M-\beta} \\ &< \sum_{\beta=\lfloor (h_{\theta,\max}+\varepsilon)M \rfloor+1}^{M} \binom{M}{\beta} \left(h_{\theta,\max}\right)^\beta \left(1-h_{\theta,\max}\right)^{M-\beta} \\ &< \varepsilon. \end{aligned} \tag{D.20}$$

Moreover, if

$$\nu_1 > \max_{\theta\in\Theta\setminus\{1\}} \left(h_{\theta,\max}+\varepsilon\right) = \max_{\theta\in\Theta\setminus\{1\},k\in\Theta} \left(q(\theta,k)+\varepsilon\right),$$

then

$$\overline{p}^{\theta 1} < \varepsilon, \qquad \forall \theta \in \Theta \setminus \{1\}. \tag{D.21}$$

Second, one derives the lower bound of $\overline{p}^{11}$. Denote $\zeta = \min_{\theta\in\Theta\setminus\{1\}} q(1,\theta)$, one has

$$\begin{aligned} h_1\left(\overline{\mathbf{p}}^1\right) &= \frac{p_1 q(1,1)\,\overline{p}^{11} + \sum_{k\in\Theta\setminus\{1\}} p_k q(1,k)\,\overline{p}^{k1}}{\sum_{k\in\Theta} p_k \overline{p}^{k1}} \\ &\geqslant \frac{p_1 q(1,1)\,\overline{p}^{11} + \zeta \sum_{k\in\Theta\setminus\{1\}} p_k \overline{p}^{k1}}{\sum_{k\in\Theta} p_k \overline{p}^{k1}} \\ &> \frac{p_1 q(1,1)\,p^{11}_{\min}(1) + \zeta\varepsilon \sum_{k\in\Theta\setminus\{1\}} p_k}{p_1 p^{11}_{\min}(1) + \varepsilon \sum_{k\in\Theta\setminus\{1\}} p_k} = \chi. \end{aligned}$$

According to Lemma D.2, one has $\overline{p}^{11} > g(\chi)$. If $\nu_1 < \chi - \varepsilon$, $\overline{p}^{11}$ can be further bounded as

$$\begin{aligned} \overline{p}^{11} > g(\chi) &= \sum_{\beta=\lceil \nu_1 M\rceil}^{M} \binom{M}{\beta} \chi^{\beta} (1-\chi)^{M-\beta} \\ &> \sum_{\beta=\lfloor (\chi-\varepsilon)M\rfloor+1}^{M} \binom{M}{\beta} \chi^{\beta} (1-\chi)^{M-\beta} \\ &\geqslant \sum_{\beta=\lfloor (\chi-\varepsilon)M\rfloor+1}^{\lceil (\chi+\varepsilon)M\rceil-1} \binom{M}{\beta} \chi^{\beta} (1-\chi)^{M-\beta} \\ &\overset{(a)}{>} 1-\varepsilon. \end{aligned}$$

Again, similar to (D.4), there exists $m_2 > m_1$ such that for all $M > m_2$, (a) is true.

Notice that $\lim_{\varepsilon\to 0} \chi = q(1,1)$, one obtains that if

$$\max_{\theta\in\Theta\setminus\{1\}, k\in\Theta} q(\theta,k) < \nu_1 < q(1,1),$$

then $\lim_{M\to\infty} \overline{p}^{11} = 1$. Using similar derivations, if $q(\theta+1,1) < \nu_\theta < q(\theta,1)$, for all $\theta = 1,\ldots,K-1$, one has $\lim_{M\to\infty} \overline{p}^{\theta\theta} = 1$, $\forall\theta \in \Theta\setminus\{1\}$.

D.4 Proof of Proposition 5.6

One starts with Case I. From (5.27), one has

$$\begin{aligned} \sum_{\theta\in\Theta} \overline{p}^{\theta\theta} &= \sum_{\theta=1}^{K} \sum_{\beta:\nu_\theta \leqslant \beta/M < \nu_{\theta-1}} \overline{X}_{\theta}^{M,\beta} \\ &= \sum_{\beta=\lceil \nu_1 M\rceil}^{M} \binom{M}{\beta} (s_1)^{\beta} (1-s_1)^{M-\beta} + \sum_{\theta=2}^{K} \sum_{\beta=\lceil \nu_\theta M\rceil}^{\lceil \nu_{\theta-1} M\rceil-1} \binom{M}{\beta} (s_\theta)^{\beta} (1-s_\theta)^{M-\beta} \end{aligned} \tag{D.22}$$

To lighten the notations, let $d_\theta = \lceil \nu_\theta M \rceil$. Then one needs to find $\mathbf{d}^* = (d_1^* \dots d_{K-1}^*)$ such that $d_1^* > \dots > d_{K-1}^*$ and

$$U(\mathbf{d}) = \sum_{d=d_1}^{M} \binom{M}{\beta} (s_1)^\beta (1-s_1)^{M-\beta} + \sum_{\theta=2}^{K} \sum_{d=d_\theta}^{d_\theta - 1} \binom{M}{\beta} (s_\theta)^\beta (1-s_\theta)^{M-\beta} \tag{D.23}$$

takes maximum value as $\mathbf{d} = \mathbf{d}^*$.

Consider an arbitrary $\theta \in \Theta$, one evaluates

$$\begin{aligned} \delta_\theta &= U(d_1 \dots, d_\theta + 1, \dots d_{K-1}) - U(d_1 \dots, d_\theta, \dots d_{K-1}) \\ &= \binom{M}{d_\theta} \left((s_{\theta+1})^{d_\theta} (1-s_{\theta+1})^{M-d_\theta} - (s_\theta)^{d_\theta} (1-s_\theta)^{M-d_\theta} \right). \end{aligned} \tag{D.24}$$

If $\delta_\theta > 0$, then

$$(s_{\theta+1})^{d_\theta} (1-s_{\theta+1})^{M-d_\theta} > (s_\theta)^{d_\theta} (1-s_\theta)^{M-d_\theta}, \tag{D.25}$$

leading to

$$d_\theta < M \frac{\log\left(\frac{1-s_\theta}{1-s_{\theta+1}}\right)}{\log\left(\frac{s_{\theta+1}}{s_\theta}\frac{1-s_\theta}{1-s_{\theta+1}}\right)} = \widetilde{d}_\theta, \tag{D.26}$$

notice that $\log\left(\frac{s_{\theta+1}}{s_\theta}\frac{1-s_\theta}{1-s_{\theta+1}}\right) < 0$ as $s_{\theta+1} < s_\theta$.

In contrast, if $\delta_\theta < 0$, then $d_\theta > \widetilde{d}_\theta$. As a conclusion, for any $d_1 > \dots d_\theta \dots > d_{K-1}$, one has

$$U(d_1 \dots, d_\theta, \dots d_K) \leqslant U(d_1 \dots, d_\theta^*, \dots d_K), \quad \forall \theta = 1 \dots (K-1) \tag{D.27}$$

where

$$d_\theta^* = \left\lceil \widetilde{d}_\theta \right\rceil = \left\lceil M \frac{\log\left(\frac{1-s_\theta}{1-s_{\theta+1}}\right)}{\log\left(\frac{s_{\theta+1}}{s_\theta}\frac{1-s_\theta}{1-s_{\theta+1}}\right)} \right\rceil. \tag{D.28}$$

Then one has

$$\begin{aligned} U(d_1 \dots d_{K-1}) &\leqslant U(d_1^*, d_2 \dots d_{K-1}) \leqslant U(d_1^*, d_2^*, d_3 \dots d_{K-1}) \\ &\leqslant U(d_1^*, d_2^*, d_3 \dots d_{K-1}) \leqslant \cdots \\ &\leqslant U(d_1^* \dots d_K^*). \end{aligned} \tag{D.29}$$

Replacing s_θ with $q(\theta, 1)$ and using the same derivations, one obtains the ν_θs which maximize $\sum_{\theta \in \Theta} \widehat{p}^{\theta\theta}$.

Bibliography

[AB10] J. Audibert and S. Bubeck. "Best arm identification in multi-armed bandits". In: *COLT-23th Conf. on Learning Theory*. 2010.

[AF12] E. Ayday and F. Fekri. "An iterative algorithm for trust management and adversary detection for delay-tolerant networks". In: *IEEE Trans. on Mobile Computing* 11.9 (2012), pp. 1514–1531.

[Ahl+00] R. Ahlswede et al. "Network information flow". In: *IEEE Trans. Inf. Theory* 46 (2000), pp. 1204–1216.

[Ang+09] H.H. Ang et al. "Communication-efficient classification in P2P networks". In: *Proc. Joint European Conf. on Machine Learning and Knowledge Discovery in Databases*. Springer. 2009, pp. 83–98.

[Ban22] S. Banach. "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales". In: *Fund. Math.* 3.1 (1922), pp. 133–181.

[Bas+12] F. Bassi et al. "Compressive linear network coding for efficient data collection in wireless sensor networks". In: *Proc. 20th European Signal Processing Conf. (EUSIPCO)*. Bucharest, Romania, 2012, pp. 714 –718.

[BBC90] G. Belforte, B. Bona, and V. Cerone. "Parameter estimation algorithms for a set-membership description of uncertainty". In: *Automatica* 26.5 (1990), pp. 887–898.

[BL94] V. Barnett and T. Lewis. *Outliers in statistical data*. Vol. 3. Wiley New York, 1994.

[BM11] M. Bayati and A. Montanari. "The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing". In: *IEEE Trans. Inf. Theory* 57.2 (2011), pp. 764–785.

[Bra+13] J. W. Branch et al. "In-network outlier detection in wireless sensor networks". In: *Knowledge and information systems* 34.1 (2013), pp. 23–54.

[BSB10] D. Baron, S. Sarvotham, and R. G. Baraniuk. "Bayesian Compressive Sensing Via Belief Propagation". In: *IEEE Trans. Signal Process.* 58.1 (2010), pp. 269 –280.

[BTF12] I. Bourtsoulatze, N. Thomos, and P. Frossard. "Correlation-Aware Reconstruction of Network Coded Sources". In: *Proc. IEEE International Symposium on Network Coding (NetCod)*. Cambridge, MA. 2012, pp. 91–96.

[Che+11] M. Cheraghchi et al. "Group testing with probabilistic tests: theory, design and application". In: *IEEE Trans. Inf. Theory* 57.10 (2011), pp. 7057 –7067.

[Chi+11] A. Chiuso et al. "Gossip algorithms for simultaneous distributed estimation and classification in sensor networks". In: *IEEE Journal of Selected Topics in Signal Processing* 5.4 (2011), pp. 691–706.

[Cho+09] J.-Y. Choi et al. "A distributed adaptive scheme for detecting faults in wireless sensor networks". In: *WSEAS Transactions on Communications* 8.2 (2009), pp. 269–278.

[CKS06] J. Chen, S. Kher, and A. Somani. "Distributed fault detection of wireless sensor networks". In: *Proc Workshop DIWANS*. New York, NY, 2006, pp. 65 –72.

[CRT06] E. Candes, J. Romberg, and T. Tao. "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information". In: *IEEE Trans. Inf. Theory* 52.2 (2006), pp. 489 –509.

[Csi82] I. Csiszar. "Linear codes for sources and source networks: Error exponents, universal coding". In: *IEEE Trans. Inf. Theory* 28.4 (1982), pp. 585–592.

[Csi98] I. Csiszar. "The Method of Types". In: *IEEE Trans. Inf. Theory* 44.6 (1998), pp. 2505–2523.

[CT05] E. Candes and T. Tao. "Decoding by linear programming". In: *IEEE Trans. Inf. Theory* 51.12 (2005), pp. 4203 –4215.

[CT06a] E. Candes and T. Tao. "Near optimal signal recovery from random projections: Universal encoding strategies?" In: *IEEE Trans. Inf. Theory* 52 (2006), pp. 5406 –5425.

[CT06b] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Vol. 2. Wiley-Interscience, 2006.

[CT12] T.M. Cover and J.A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[CWJ03] P. Chou, Y. Wu, and K. Jain. "Practical network coding". In: *Proc. 41-st Allerton Conf.* Monticello, IL (2003).

[DH93] D.-Z. Du and F.K. Hwang. *Combinatorial group testing and its applications*. World Scientific, 1993.

[Dim+10] A.G. Dimakis et al. "Gossip algorithms for distributed signal processing". In: *Proceedings of the IEEE* 98.11 (2010), pp. 1847–1864.

[Din+05] M. Ding et al. "Localized fault-tolerant event boundary detection in sensor networks". In: *Proc IEEE INFOCOM*. Miami, FL, 2005, pp. 902 –913.

[DM09] S.C. Draper and S. Malekpour. "Compressed sensing over finite fields". In: *Proc. Intl. Symp. Inf. Theory*. Seoul, Korea, 2009, pp. 669 –673.

[Don+14] H. Dong et al. "A Survey on Distributed Filtering and Fault Detection for Sensor Networks". In: *Mathematical Problems in Engineering* (2014).

[Don06] D. Donoho. "Compressed sensing". In: *IEEE Trans. Inf. Theory* 52 (2006), pp. 1289–1306.

[Dou02] J.R. Douceur. "The sybil attack". In: *Peer-to-peer Systems*. Springer, 2002, pp. 251–260.

[DP+15] A. De Paola et al. "Adaptive Distributed Outlier Detection for WSNs". In: *Cybernetics, IEEE Transactions on* 45.5 (2015), pp. 888–899. ISSN: 2168-2267.

[EP06] N. Eagle and A. Pentland. "Reality mining: sensing complex social systems". In: *Personal and ubiquitous computing* 10.4 (2006), pp. 255–268.

[Fal03] K. Fall. "A delay-tolerant network architecture for challenged internets". In: *Proc. Conf. on Applications, technologies, architectures, and protocols for computer communications*. ACM. 2003, pp. 27–34.

[FFR14] F. Fagnani, S. Fosson, and C. Ravazzi. "A distributed classification/estimation algorithm for sensor networks". In: *SIAM Journal on Control and Optimization* 52.1 (2014), pp. 189–218.

[FM11] S. Feizi and M. Médard. "A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing". In: *Proc. 49th Annual Allerton Conf.* IEEE. 2011, pp. 1048–1054.

[FZ08] F. Fagnani and S. Zampieri. "Randomized consensus algorithms over large scale networks". In: *IEEE Journal on Selected Areas in Communications* 26.4 (2008), pp. 634–649.

[Gal68] R.G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Sons., 1968.

[GD13] A. Granas and J. Dugundji. *Fixed point theory*. Springer Science & Business Media, 2013.

[GLG16] L. Galluccio, B. Lorenzo, and S. Glisic. "Sociality-aided new adaptive infection recovery schemes for multicast DTNs". In: *IEEE Trans. on Vehicular Technology* 65.5 (2016), pp. 3360–3375.

[Guo+15] B. Guo et al. "Mobile crowd sensing and computing: the review of an emerging human-powered sensing paradigm". In: *ACM Computing Surveys* 48.1 (2015).

[GXL07] J.-L. Gao, Y.-J. Xu, and X.-W. Li. "Weighted-median based distributed fault detection for wireless sensor networks". In: *Journal of Software* 18.5 (2007), pp. 1208 –1217.

[Hae+09] M. Haenggi et al. "Stochastic geometry and random graphs for the analysis and design of wireless networks". In: *IEEE Journal on Selected Areas in Communications* 27.7 (2009), pp. 1029–1046.

[Hae12] M. Haenggi. *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[Han+13] Yu Han et al. "A reputation-aware decision-making approach for improving the efficiency of crowdsourcing systems". In: *Proc. International Conf. on Autonomous agents and Multi-agent Systems*. Richland, SC, 2013, pp. 1315–1316.

[Han+15] G. Han et al. "An Attack-Resistant Trust Model Based on Multidimensional Trust Metrics in Underwater Acoustic Sensor Network". In: *IEEE Trans. on Mobile Computing* 14.12 (2015), pp. 2447–2459.

[HCY11] P. Hui, J. Crowcroft, and E. Yoneki. "Bubble rap: Social-based forwarding in delay-tolerant networks". In: *IEEE Trans. on Mobile Computing* 10.11 (2011), pp. 1576–1589.

[Hec+16] R. Heckel et al. "Active Ranking from Pairwise Comparisons and the Futility of Parametric Assumptions". In: *arXiv preprint arXiv:1606.08842* (2016).

[HM09] M. Huang and J.H. Manton. "Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior". In: *SIAM Journal on Control and Optimization* 48.1 (2009), pp. 134–161.

[HN06] J. Haupt and R. Nowak. "Signal Reconstruction From Noisy Random Projections". In: *IEEE Trans. Inf. Theory* 52.9 (2006), pp. 4036 –4048.

[Ho+06] T. Ho et al. "A random linear network coding approach to multicast". In: *IEEE Trans. Inf. Theory* 52 (2006), pp. 4413–4430.

[HO+15] E. Hernandez-Orallo et al. "CoCoWa: A Collaborative Contact-based Watchdog for Detecting Selfish Nodes". In: *IEEE Trans. on Mobile Computing* 14.6 (2015), pp. 1162–1175.

[IKAA12] L. Iwaza, M. Kieffer, and K. Al-Agha. "MAP estimation of network-coded correlated sources". In: (2012), pp. 199–202.

[Ins] Texas Instruments. *CC253x/4x User's Guide*. http://www.ti.com/lit/ug/swru191f/swru191f.pdf.

[Ise05] R. Isermann. "Model-based fault-detection and diagnosis–status and applications". In: *Annual Reviews in control* 29.1 (2005), pp. 71–85.

[Jaf+09] M. Jafari et al. "Compressed network coding vectors". In: *Proc. IEEE Intl. Symp. on Inf. Theory*. Seoul, Korea, 2009, pp. 109–113.

[Jau+01] L. Jaulin et al. *Applied Interval Analysis*. London: Springer-Verlag, 2001.

[Ji+10] S. Ji et al. "Distributed fault detection for wireless sensor based on weighted average". In: *Proc NSWCTC*. Wuhan, China, 2010, pp. 57 –60.

[JXC08] S. Ji, Y. Xue, and L. Carin. "Bayesian Compressive Sensing". In: *IEEE Trans. Signal Process.* 52.6 (2008), pp. 2346 –2356.

[KAF12] M.J. Khabbaz, C.M. Assi, and W.F. Fawaz. "Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges". In: *IEEE Communications Surveys & Tutorials* 14.2 (2012), pp. 607–640.

[KFL01] F.R. Kschischang, B.J. Frey, and H.A. Loeliger. "Factor graphs and the sum-product algorithm". In: *IEEE Trans. Inf. Theory* 47.2 (2001), pp. 498–519.

[KM14a] B. Kantarci and H. T. Mouftah. "Trustworthy Sensing for Public Safety in Cloud-Centric Internet of Things". In: *IEEE Internet Of Things Journal* 1.4 (2014).

[KM14b] B. Kantarci and H.T. Mouftah. "Reputation-based sensing-as-a-service for crowd management over the cloud". In: *Proc. of IEEE Int. Conf. on Communications, ICC*. 2014.

[KSA12] MHR Khouzani, S. Sarkar, and E. Altman. "Maximum damage malware attack in mobile wireless networks". In: *IEEE/ACM Transactions on Networking* 20.5 (2012), pp. 1347–1360.

[LBK14] W. Li, F. Bassi, and M. Kieffer. "Robust Bayesian compressed sensing over finite fields: asymptotic performance analysis". In: *arXiv preprint arXiv:1401.4313* (2014).

[LBK16] W. Li, F. Bassi, and M. Kieffer. "Sparse Random Linear Network Coding for Data Compression in WSNs". In: *2016 IEEE International Symposium on Information Theory (ISIT)*. 2016.

[LC08] M.-H. Lee and Y.-H. Choi. "Fault detection of wireless sensor networks". In: *Computer Communications* 31.14 (2008), pp. 3469–3475.

[LCJ15] M. Leinonen, Ma. Codreanu, and M. Juntti. "Sequential Compressed Sensing With Progressive Signal Reconstruction in Wireless Sensor Networks". In: *Wireless Communications, IEEE Transactions on* 14.3 (2015), pp. 1622–1635.

[Li+15a] W. Li et al. "Iterative distributed outlier detection for wireless sensor networks: Equilibrium and convergence analysis". In: *Proc. IEEE Conf. on Decision and Control (CDC)*. IEEE. 2015, pp. 3050–3056.

[Li+15b] W. Li et al. "Low-complexity distributed fault detection for wireless sensor networks". In: *Proc. IEEE ICC*. London, United Kingdom, June 2015.

[Li+16a] W. Li et al. "Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests". In: *IEEE Trans. on Signal and Information Processing over Networks* 2.1 (2016), pp. 29–48.

[Li+16b] W. Li et al. "Distributed Faulty Node Detection in DTNs". In: *Proc. International Conf. on Computer Communication and Networks*. 2016.

[Li+16c] W. Li et al. "Impact of channel access issues and packet losses on distributed outlier detection within wireless sensor networks". In: *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. 2016, pp. 3746–3750.

[Li+16d] W. Li et al. "Distributed Faulty Node Detection in Delay Tolerant Networks: Design and Analysis". In: (2016, hal-01327472).

[Li+16e] W. Li et al. "Self-Rating in a Community of Peers". In: (2016, hal-01327792).

[Li+16f] W. Li et al. "Peer-Assisted Individual Assessment in a Multi-Agent System". In: (2016, hal-01349638).

[Liu+15] Y. Liu et al. "The Mason test: A defense against Sybil attacks in wireless networks without trusted authorities". In: *IEEE Trans. on Mobile Computing* 14.11 (2015), pp. 2376–2391.

[Lo+13] C. Lo et al. "Efficient sensor fault detection using combinatorial group testing". In: *IEEE Int. Conf. on DCOSS*. 2013, pp. 199–206.

[Luo+07] P. Luo et al. "Distributed classification in peer-to-peer networks". In: *Proc. 13th ACM SIGKDD International Conf. on Knowledge Discovery and Data mining*. ACM. 2007, pp. 968–976.

[Mil+96] M. Milanese et al., eds. *Bounding Approaches to System Identification*. New York, NY: Plenum Press, 1996.

[MK13] A. Mahapatro and P.M. Khilar. "Fault diagnosis in wireless sensor networks: A survey". In: *IEEE Communications Surveys & Tutorials* 15.4 (2013), pp. 2000–2026.

[Mon12] A. Montanari. "Graphical models concepts in compressed sensing". In: *Compressed Sensing: Theory and Applications*. 2012, pp. 394–438.

[NL14] M. Nabaee and F. Labeau. "Quantized Network Coding for correlated sources". In: *EURASIP Journal on Wireless Communications and Networking* 2014.1 (2014), pp. 1–17.

[Orl] M. Orlinski. *Encounter traces for the ONE simulator*. Downloaded from "http://www.shigs.co.uk/index.php?page=traces".

[Pan+15] M. Panda et al. "Tracking Message Spread in Mobile Delay Tolerant Networks". In: *IEEE Trans. on Mobile Computing* 14.8 (2015), pp. 1737–1750.

[Pen+14] W. Peng et al. "Behavioral Malware Detection in Delay Tolerant Networks". In: *IEEE Trans. on Parallel and Distributed Systems* 25.1 (2014), pp. 53–63.

[PK15] M. Panda and PM Khilar. "Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test". In: *Ad Hoc Networks* 25 (2015), pp. 170–184.

[RAG12] K. Rajawat, C. Alfonso, and G.B. Giannakis. "Network-Compressive Coding for Wireless Sensors with Correlated Data". In: *Wireless Communications, IEEE Trans. on* 11.12 (2012), pp. 4264–4274.

[Ren+15] J. Ren et al. "SACRM: Social Aware Crowdsourcing with Reputation Management in mobile sensing". In: *Computer Communications* 65 (2015).

[Rud64] W. Rudin. *Principles of mathematical analysis*. Vol. 3. McGraw-Hill New York, 1964.

[SB98] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. Vol. 1. MIT press Cambridge, 1998.

[SBW16] Nihar B Shah, Sivaraman Balakrishnan, and Martin J Wainwright. "A Permutation-based Model for Crowd Labeling: Optimal Estimation and Robustness". In: *arXiv preprint arXiv:1606.09632* (2016).

[Sco+09] J. Scott et al. *CRAWDAD dataset cambridge/haggle (v. 2009-05-29)*. Downloaded from http://crawdad.org/cambridge/haggle/20090529. May 2009. DOI: 10.15783/C70011.

[She+13] X. Sheng et al. "Sensing as a Service: challenges, solutions and future directions". In: *IEEE Sensors Journal* 13 (10 2013).

[SL13] J-T. Seong and H-N. Lee. "Necessary and sufficient conditions for recovery of sparse signals over finite fields". In: *IEEE Communications Letters* 17.10 (2013), pp. 1976 – 1979.

[SRF14] V. NGJ Soares, J. JPC Rodrigues, and F. Farahmand. "GeoSpray: A geographic routing protocol for vehicular delay-tolerant networks". In: *Information Fusion* 15 (2014), pp. 102–113.

[TBD12] V.Y.F. Tan, L. Balzano, and S.C. Draper. "Rank minimization over finite fields: fundamental limits and coding-theoretic interpretations". In: *IEEE Trans. Inf. Theory* 58.4 (2012), pp. 2018–2039.

[Tim] *TIMAC Software*. http://www.ti.com/timac.

[TTF13] T. Tosic, N. Thomos, and P. Frossard. "Distributed sensor failure detection in sensor networks". In: *Signal Processing* 93.2 (2013), pp. 399–410.

[Ver+10] R. Verdone et al. *Wireless sensor and actuator networks: technologies, analysis and design*. Academic Press, 2010.

[Yeu04] R.W. Yeung. "A first course in information theory". In: *MA: Kluwer* Princeton, NJ (2004).

[YS12] Z. Yu and M. Van der Schaar. "Reputation-based incentive protocols in crowdsourcing applications". In: *Proc. IEEE INFOCOM*. Orlando, FL, 2012, pp. 2140–2148.

[Zhu+10] H. Zhu et al. "Recognizing Exponential Inter-Contact Time in VANETs". In: *Proc. INFOCOM*. 2010, pp. 1–5.

[Zhu+12] B. Zhu et al. "Distributed faulty node detection and isolation in delay-tolerant vehicular sensor networks". In: *Proc. PIMRC*. 2012, pp. 1497–1502.

[Zhu+14] H. Zhu et al. "A probabilistic misbehavior detection scheme toward efficient trust establishment in delay-tolerant networks". In: *IEEE Trans. on Parallel and Distributed Systems* 25.1 (2014), pp. 22–32.

[ZMH10] Y. Zhang, N. Meratnia, and P. Havinga. "Outlier detection techniques for wireless sensor networks: A survey". In: *IEEE Communications Surveys & Tutorials* 12.2 (2010), pp. 159–170.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Collecte et estimation robustes d'information dans un réseau de capteurs sans fils

Mots clés : compression distribuée, estimation distribuée, équilibre, réseau de capteurs sans fils, réseau tolérant aux déconnections

Résumé : Les réseaux de capteurs sans fils (RCSFs) suscitent un intérêt croissant depuis une vingtaine d'années. La première partie de cette thèse est consacré à l'étude de l'efficacité de compression de données corrélées provenant d'un RCSF et acheminées vers un point de collecte à l'aide du codage réseau linéaire aléatoire. Les conditions nécessaires et suffisantes sont obtenues pour récupérer parfaitement les données que les capteurs mesurent. Puis on considère les nœuds dans un RCSF collaborant afin d'exécuter une tâche donnée (acquisition, détection...), pour laquelle chaque nœud a potentiellement un niveau d'expertise différent. La seconde partie de cette thèse est dédiée à la conception et à l'analyse d'algorithmes d'auto-évaluation distribués (AED), qui permettent à chaque nœud d'auto-évaluer son niveau d'expert. Trois types de problèmes sont considérés: *i*) la détection distribuée des nœuds défaillants (DDD), qui permet d'identifier les nœuds équipés de capteurs défectueux dans un RCSF; *ii*) la DDD dans un réseau tolérant aux déconnections (RTD) dont la topologie est dynamique et le degré de connectivité très faible; *iii*) la AED avec interactions pair à pair. Les résultats théoriques sont utiles pour configurer les paramètres des algorithmes.

Title : Distributed Information Gathering and Estimation in Wireless Sensor Networks

Keywords : distributed compression, distributed estimation, equilibrium, wireless sensor network, delay tolerant network

Abstract : Wireless sensor networks (WSNs) have attracted much interests in the last decade. The first part of this thesis considers sparse random linear network coding is for data gathering and compression in WSNs. An information-theoretic approach is applied to demonstrate the necessary and sufficient conditions to realize the asymptotically perfect reconstruction under MAP estimation. The second part of the thesis concerns the distributed self-rating (DSR) problem, for WSNs with nodes that have different ability of performing some task (sensing, detection...). The main assumption is that each node does not know and needs to estimate its ability. Depending on the number of ability levels and the communication conditions, three sub-problems have been addressed: *i)* distributed faulty node detection (DFD) to identify the nodes equipped with defective sensors in dense WSNs; *ii*) DFD in delay tolerant networks (DTNs) with sparse and intermittent connectivity; *iii*) DSR using pairwise comparison. Distributed algorithms have been proposed and analyzed. Theoretical results assess the effectiveness of the proposed solution and give guidelines in the design of the algorithm.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France