

Estimation Distribuée basée sur un Consensus en Temps Fini

Alain Y. KIBANGOU

GIPSA-Lab, Université Joseph Fourier, CNRS
961 rue de la Houille Blanche, BP 46, F- 38402 Grenoble Cedex, France
alain.kibangou@ujf-grenoble.fr

Résumé – Dans cet article, nous traitons le problème de l'estimation distribuée dans un réseau de capteurs. En vue d'atteindre des performances similaires à celles obtenues dans un contexte centralisé, nous proposons un algorithme de consensus en temps fini. Ce consensus est atteint grâce à une séquence de matrices de consensus conjointement diagonalisables dont la synthèse est faite à partir de la matrice Laplacienne du graphe représentant les interactions entre nœuds du réseau.

Abstract – In this paper, we address the problem of distributed estimation with a sensor network. In order to achieve performance similar to those obtained in a centralized framework, we propose a consensus algorithm in finite time. This consensus is achieved owing to a sequence of jointly diagonalizable consensus matrices derived from the graph Laplacian matrix.

1 Introduction

Depuis quelques années, en raison de leur grand potentiel en termes d'applications, les réseaux de capteurs, constitués de dispositifs autonomes et communicants, font l'objet d'une attention particulière et d'un grand nombre d'études. Ils posent, en effet, de nouveaux problèmes scientifiques et techniques dont la résolution s'avère impérieuse. Dans cet article, nous nous intéressons à la problématique de l'estimation, inhérente à la fusion des données. Les différents nœuds-capteurs étant dotés de suffisamment de puissance de calcul et de communication, il s'agit ici de réaliser l'estimation d'un vecteur de paramètres non pas en fusionnant les données en un nœud central mais en distribuant le calcul sur tout le réseau [1]. Dans la littérature, les solutions sont généralement obtenues via un consensus entre les différents nœuds-capteurs du réseau, consensus portant généralement sur une valeur moyenne. Notons qu'en plus de la fusion des données [2], les problèmes de consensus apparaissent aussi dans des domaines comme le contrôle de formation des véhicules [3, 4], le calcul distribué [5], le traitement du signal tensoriel distribué [6], et dans bien d'autres domaines d'applications (Voir [7] et [8] ainsi que les références qui y sont citées).

Le consensus sur la moyenne peut être obtenu en permettant à chaque nœud de mettre à jour, de manière itérative, sa propre valeur comme une combinaison linéaire de cette dernière et de celles reçues de la part de ses voisins. En choisissant de manière adéquate les pondérations utilisées dans ce processus itératif, on montre que tous les nœuds peuvent converger vers une même valeur, à savoir la valeur moyenne des valeurs initiales des nœuds du réseau. Dans la plupart des contributions de la littérature, seules des propriétés de convergence asymptotique sont prouvées. Cependant, dans plusieurs cas, la

valeur moyenne sert à alimenter des processus d'estimation, de détection ou de commande qui sont dynamiques [2, 6]. Il est évident qu'une convergence asymptotique n'est pas souhaitable dans ce cas. En effet, on est tenu d'arrêter l'algorithme du consensus au bout d'un certain nombre d'itérations, ce qui peut être source d'erreur pour le processus utilisateur de la valeur moyenne estimée. Certains auteurs se sont alors focalisés sur l'accélération de la convergence de ces algorithmes [9, 10]. D'autres, très peu, ont étudié la question de la convergence en temps fini de ces algorithmes. Dans un formalisme lié aux systèmes à temps discret, la question a été brièvement évoquée dans [11]. Cependant, la méthode requiert un graphe complètement connecté pendant au moins un créneau de temps. Une autre approche consiste à calculer les pondérations de sorte que la matrice d'observabilité du système appartienne à un certain espace. L'idée de base étant qu'au bout d'un certain temps, les nœuds ont récolté suffisamment d'observations pour reconstruire l'état initial du système [12]. Dès lors, chaque nœud peut alors calculer la valeur exacte de la moyenne. Les mêmes auteurs ont aussi montré que chaque nœud peut calculer la valeur du consensus comme une combinaison linéaire de ses propres valeurs passées au bout de D itérations, D étant le degré du polynôme minimal associé à la matrice des pondérations [13]. Dans [14], en se basant sur les propriétés des graphes de de Bruijn et du produit de Kronecker par bloc, il a été montré que le consensus sur la moyenne peut être obtenu en temps fini si le nombre de nœuds est exactement une puissance du degré maximal du graphe. D'autres études intéressantes concernent le contexte pair-à-pair [15] et la factorisation de matrices [16]. Cette dernière contribution requiert hélas une architecture centralisée. Dans cet article nous montrons que, dans un contexte distribué, la convergence en temps fini peut être obtenue en utilisant des matrices conjointement diagonalisables, formées

à partir de la matrice Laplacienne du graphe. Cette approche requiert toutefois une étape de configuration du réseau durant laquelle la séquence de matrices est générée.

2 Position du problème

Considérons l'estimation d'un vecteur de paramètres inconnus $\boldsymbol{\theta} \in \mathfrak{R}^M$ par N nœuds-capteurs spatialement distribués. Chaque un de ces nœuds-capteurs réalise une mesure bruitée de ce vecteur de paramètres au travers d'un modèle linéaire local caractérisé par une matrice $\mathbf{A}_i \in \mathfrak{R}^{M_i \times M}$:

$$\mathbf{y}_i = \mathbf{A}_i \boldsymbol{\theta} + \mathbf{v}_i, \quad i = 1, \dots, N$$

Le bruit \mathbf{v}_i est supposé de moyenne nulle, de matrice de covariance \mathbf{S}_i , et indépendant d'un capteur à l'autre.

L'estimé de $\boldsymbol{\theta}$ au sens du maximum de vraisemblance est alors donné par :

$$\hat{\boldsymbol{\theta}} = \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{y}_i \right). \quad (1)$$

Définissons à présent les quantités locales suivantes : $\mathbf{P}_i(0) = \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{A}_i$ et $\mathbf{q}_i(0) = \mathbf{A}_i^T \mathbf{S}_i^{-1} \mathbf{y}_i$. L'objectif de l'algorithme du consensus est de faire converger toutes les valeurs locales vers la même valeur en occurrence la valeur moyenne des valeurs initiales, i.e.

$$\lim_{k \rightarrow \infty} \mathbf{P}_i(k) = \mathbf{P} = \frac{1}{N} \sum_{i=1}^N \mathbf{P}_i(0)$$

et

$$\lim_{k \rightarrow \infty} \mathbf{q}_i(k) = \mathbf{q} = \frac{1}{N} \sum_{i=1}^N \mathbf{q}_i(0).$$

Dès lors, chaque nœud peut alors estimer le vecteur des paramètres inconnus comme :

$$\hat{\boldsymbol{\theta}}_i = \mathbf{P}^{-1} \mathbf{q}.$$

Ce qui équivaut à l'estimé qu'on aurait obtenu si toutes les observations des capteurs étaient récoltées en un nœud central.

Soient les vecteurs locaux $\mathbf{x}_i = \text{vec}(\mathbf{P}_i)$ ou $\mathbf{x}_i = \mathbf{q}_i$, l'opérateur $\text{vec}(\cdot)$ concaténant les colonnes de la matrice en argument de sorte à former un seul vecteur colonne. Définissons alors la matrice $\mathbf{X} = (\mathbf{x}_1 \ \dots \ \mathbf{x}_N)^T$. Pour obtenir le consensus sur la moyenne dans l'un ou l'autre cas, l'algorithme du consensus consiste à itérer la relation suivante :

$$\mathbf{X}(k+1) = \mathbf{W}\mathbf{X}(k) \quad (2)$$

la matrice \mathbf{W} étant consistante avec le réseau ; c'est-à-dire que les éléments $w_{i,j}$ de \mathbf{W} sont nuls s'il n'y a aucune interaction entre les nœuds i et j . Le consensus sur la moyenne est atteint si $\lim_{k \rightarrow \infty} \mathbf{x}(k) = \frac{1}{N} \mathbf{1} \mathbf{1}^T \mathbf{x}(0)$, signifiant que $\lim_{k \rightarrow \infty} \mathbf{W}^k = \frac{1}{N} \mathbf{1} \mathbf{1}^T$.

Il est bien connu que la condition nécessaire et suffisante pour atteindre le consensus de manière asymptotique est que la matrice de pondération \mathbf{W} ait 1 comme valeur singulière simple tandis que toutes les autres valeurs singulières sont de module

inférieur à 1, les vecteurs propres à gauche et à droite associés à la valeur propre unité étant $\frac{1}{N} \mathbf{1}$ et $\mathbf{1}$ [9].

Notre objectif est de trouver un ensemble fini de matrices $\{\mathbf{W}_i\}_{i=1, \dots, D}$, consistantes avec le réseau, telles que :

$$\prod_{i=1}^D \mathbf{W}_i = \frac{1}{N} \mathbf{1} \mathbf{1}^T. \quad (3)$$

3 Synthèse de matrices permettant d'assurer un consensus en temps fini

D'une manière générale, le problème (3) se traduit par un système d'équations multivariées dont la solution peut ne pas exister. Pour réduire le nombre de variables, nous proposons de choisir un ensemble de matrices telles qu'il existe une matrice orthogonale \mathbf{U} vérifiant la relation $\mathbf{W}_i = \mathbf{U} \mathbf{E}_i \mathbf{U}^T$, \mathbf{E}_i étant une matrice diagonale. Dans ce cas la relation (3) devient :

$$\mathbf{U} \prod_{i=1}^D \mathbf{E}_i \mathbf{U}^T = \frac{1}{N} \mathbf{1} \mathbf{1}^T.$$

Par ailleurs, si \mathbf{U} est une matrice orthonormale structurée telle que $\mathbf{U} = \left(\frac{1}{\sqrt{N}} \mathbf{1} \ \tilde{\mathbf{U}} \right)$ alors nous pouvons récrire le membre de droite de la relation précédente comme $\mathbf{U} \text{diag}(1, \dots, 0) \mathbf{U}^T$, $\text{diag}(\cdot)$ étant l'opérateur qui forme une matrice diagonale avec les éléments en argument. Notre problème se ramène alors à trouver une matrice orthonormale $\mathbf{U} = \left(\frac{1}{\sqrt{N}} \mathbf{1} \ \tilde{\mathbf{U}} \right)$ et des matrices diagonales \mathbf{E}_i , $i = 1, \dots, D$, telles que les matrices $\mathbf{W}_i = \mathbf{U} \mathbf{E}_i \mathbf{U}^T$ soient consistantes avec le graphe et $\prod_{i=1}^D \mathbf{E}_i = \text{diag}(1, \dots, 0)$.

La contrainte de consistance avec le graphe peut être assurée en utilisant certaines propriétés issues de la théorie spectrale des graphes. Soit \mathbf{A} la matrice d'adjacence du graphe, c'est-à-dire une matrice dont l'élément (i, j) est égal à 1 si i et j interagissent et zéro sinon. C'est une matrice de diagonale nulle. Soit N_i le nombre de nœud interagissant avec le nœud i . On définit la matrice Laplacienne \mathbf{L} du graphe comme suit : $\mathbf{L} = \mathbf{D} - \mathbf{A}$, avec $\mathbf{D} = \text{diag}(N_1, \dots, N_N)$. Si le graphe est non orienté, alors \mathbf{L} est symétrique et admet $\mathbf{1}$ comme vecteur propre associé à la valeur propre nulle. Toutes les autres valeurs propres sont strictement positives.

Dans la littérature inhérente aux problèmes de consensus, une matrice très utilisée est la matrice de Perron définie par $\mathbf{W} = \mathbf{I} - \gamma \mathbf{L}$. Elle assure la convergence asymptotique du consensus si $0 < \gamma < 1/N_{max}$, avec $N_{max} = \max\{N_1, \dots, N_N\}$. Cette matrice est consistante avec le graphe par sa dépendance directe avec la matrice Laplacienne du graphe. Afin d'assurer cette contrainte dans notre problème, nous paramétrons la séquence de matrices de la manière suivante :

$$\mathbf{W}_i = \alpha_i \mathbf{I} - \beta \mathbf{L}. \quad (4)$$

Nous obtenons par conséquent une famille de matrices conjointement diagonalisables paramétrée par les scalaires β et α_i , $i =$

$1, \dots, D$. En effet, en notant \mathbf{U} la matrice des vecteurs propres de \mathbf{L} , en obtient $\mathbf{U}^T \mathbf{W}_i \mathbf{U} = \mathbf{E}_i = \alpha_i \mathbf{I} - \beta \mathbf{A}$ où \mathbf{A} est la matrice des valeurs propres de la matrice Laplacienne ordonnées dans le sens croissant. On peut remarquer que dans cette configuration, à chaque itération, un nœud donné pondère les informations venues de ses voisins de la même manière mais modifie le poids accordé aux informations qu'il a déjà stockées.

Nous pouvons aussi envisager une autre famille de matrices en inversant les rôles des paramètres β et α_i :

$$\mathbf{W}_i = \beta \mathbf{I} - \alpha_i \mathbf{L}. \quad (5)$$

Dans ce cas, à chaque itération, chaque nœud modifie aussi bien la pondération des informations reçues que celles déjà stockées.

Les matrices \mathbf{W}_i étant consistantes avec le graphe, la détermination des paramètres inconnus est alors réalisée en assurant la contrainte $\prod_{i=1}^D \mathbf{E}_i = \text{diag}(1, \dots, 0)$.

Proposition 1 : *Etant donné un graphe non orienté, connecté, de matrice laplacienne \mathbf{L} , l'ensemble des matrices $\mathbf{W}_i = \alpha_i \mathbf{I} - \beta \mathbf{L}, i = 1, \dots, D$, permet d'atteindre le consensus sur la moyenne en D itérations si :*

1. $D + 1$ est le nombre de valeurs propres distinctes de la matrice laplacienne \mathbf{L} ;
2. les paramètres β et α_i sont donnés par :

$$\beta = \frac{1}{\left(\prod_{j=1}^D \lambda_{j+1} \right)^{1/D}} \quad \alpha_i = \beta \lambda_{i+1}, \quad i = 1, \dots, D, \quad (6)$$

$\lambda_i, i = 2, \dots, D + 1$ étant les valeurs propres, non-nulles, distinctes de \mathbf{L} .

Preuve : Soit à résoudre $\prod_{i=1}^D \mathbf{E}_i = \text{diag}(1, \dots, 0)$ avec $\mathbf{E}_i = \alpha_i \mathbf{I} - \beta \mathbf{A}$. La jème entrée diagonale de \mathbf{E}_i est donnée par $E_{i,j} = \alpha_i - \beta \lambda_j$. Par suite le système d'équations à résoudre devient :

$$\prod_{i=1}^D E_{i,1} = \prod_{i=1}^D \alpha_i = 1 \quad (7)$$

$$\prod_{i=1}^D E_{i,j} = \prod_{i=1}^D (\alpha_i - \beta \lambda_j) = 0, \quad i = 2, \dots, N \quad (8)$$

Le système donné par la seconde équation peut être redondant si certaines valeurs propres sont multiples. Si \mathbf{L} admet D valeurs propres distinctes non nulles, nous obtenons alors D équations distinctes. En profitant de leur forme factorisée, une solution paramétrée par β est donnée par $\alpha_i = \beta \lambda_{i+1}$. Par suite, sachant que $\prod_{i=1}^D \alpha_i = 1$, nous en déduisons la valeur de β .

■ Pour la seconde configuration nous énonçons la proposition suivante dont la démonstration est strictement équivalente à celle ci-dessus.

Proposition 2 : *Etant donné un graphe non orienté, connecté, de matrice laplacienne \mathbf{L} , l'ensemble des matrices $\mathbf{W}_i = \beta \mathbf{I} - \alpha_i \mathbf{L}, i = 1, \dots, D$, permet d'atteindre le consensus sur la moyenne en D itérations si :*

1. $D + 1$ est le nombre de valeurs propres distinctes de la matrice laplacienne \mathbf{L} ;
2. les paramètres β et α_i sont donnés par :

$$\beta = 1 \quad \alpha_i = \frac{1}{\lambda_{j+1}}, \quad i = 1, \dots, D, \quad (9)$$

$\lambda_i, i = 2, \dots, D + 1$ étant les valeurs propres, non-nulles, distinctes de \mathbf{L} .

Dans ce cas, le consensus en temps fini est obtenu grâce à une séquence de matrices de Perron dont les pas sont les inverses des valeurs propres de la matrice Laplacienne du graphe.

Exemple : Considérons le réseau caractérisé par la matrice laplacienne suivante :

$$\mathbf{L} = \begin{pmatrix} 3 & -1 & 0 & 0 & -1 & -1 \\ -1 & 5 & -1 & -1 & -1 & -1 \\ 0 & -1 & 3 & -1 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 & 0 \\ -1 & -1 & -1 & -1 & 5 & -1 \\ -1 & -1 & 0 & 0 & -1 & 3 \end{pmatrix}$$

Ses valeurs propres non-nulles sont 2, 4 et 6. Par conséquent l'estimation distribuée au sens du maximum de vraisemblance peut être réalisée après $D = 3$ itérations de consensus. L'ensemble des matrices de pondération étant caractérisée par $\beta = 0.2752$, $\alpha_1 = 0.5503$, $\alpha_2 = 1.1006$, et $\alpha_3 = 1.6510$. Pour la seconde configuration, la séquence de matrices de pondération est donnée par $\mathbf{W}_1 = \mathbf{I} - 0.5\mathbf{L}$, $\mathbf{W}_2 = \mathbf{I} - 0.25\mathbf{L}$ et $\mathbf{W}_3 = \mathbf{I} - \frac{1}{6}\mathbf{L}$. On notera que l'utilisation de la matrice \mathbf{W}_1 seule ne permet pas d'atteindre le consensus asymptotique car dans ce cas le système est instable.

Dans un graphe, la vitesse de circulation de l'information peut être mesurée à l'aide du diamètre du graphe ; le diamètre d'un graphe étant la distance maximale entre tous les paires de nœuds du graphe. En d'autres termes, le diamètre nous indique le nombre de pas nécessaire pour qu'une information soit partagée par les deux nœuds les plus éloignés du graphe. Une question intéressante est alors de savoir si le nombre d'itérations de l'algorithme du consensus en temps fini que nous avons proposé correspond au diamètre du graphe.

Corollaire 1 *Les matrices de pondérations des propositions 1 et 2 permettent de réaliser le consensus en un nombre d'itérations D borné en valeurs supérieures par le nombre de nœuds moins un et en valeurs inférieures par le diamètre du graphe δ .*

$$\delta \leq D < N - 1$$

Il est exactement égal au diamètre du graphe si le graphe est à distance régulière.

Preuve : *Il est bien connu qu'un graphe connecté ayant $D + 1$ valeurs propres distinctes a un diamètre δ au plus égal à D .*

Sachant que la matrice Laplacienne associée à un graphe de N nœuds admet au plus N valeurs distinctes dont la valeur propre nulle, si le graphe est non orienté, alors : $\delta \leq D < N - 1$. Il est par ailleurs bien connu que les graphes à distance régulière ont un diamètre égal au nombre de valeurs propres distinctes.

4 Simulations

Considérons le réseau de la figure 1. Nous allons comparer l'algorithme en temps fini que nous avons proposé avec une approche de calcul classique (Local degree weights) dans laquelle les pondérations sont calculées comme $w_{i,j} = 1/\max\{N_i, N_j\}$ et avec l'algorithme rapide proposée dans [9]. Nous considé-

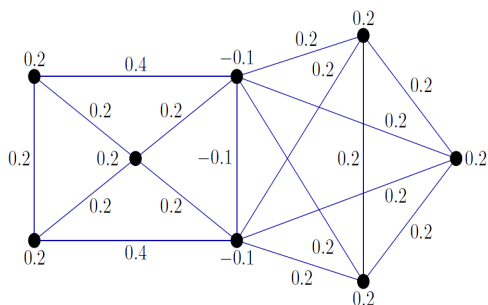


FIG. 1 – Graphe de huit nœuds avec leurs pondérations optimales au sens de [9].

rons l'estimation d'un vecteur $\theta = (-1, 4, 2)^T$ observé avec un bruit de variance 10^{-3} . Les matrices d'observation sont générées de manière aléatoire suivant une distribution gaussienne. Les résultats de la figure 2 correspondent à une moyenne sur 100 réalisations indépendantes. Nous traçons l'erreur quadratique moyenne par rapport à l'estimation dans un contexte complètement centralisé. On peut remarquer la supériorité de l'algorithme du consensus en temps fini par rapport aux algorithmes n'assurant qu'une convergence asymptotique. Il est nécessaire de signaler que les calculs intermédiaires de $\hat{\theta}$ n'ont que très peu de signification. Les valeurs significatives étant celles calculées localement (début de l'algorithme) puis celles obtenues au bout de D itérations. Ici, $D = 7$.

Il est à noter que le consensus en temps fini peut être utilisé dans une approche de type Monte-Carlo lorsque les communications sont bruitées (voir [17]). Une évaluation sur des réseaux plus large y est aussi proposée.

Références

- [1] I. Schizas, A. Ribeiro, and G. Giannakis. Consensus in ad hoc WSNs with noisy links-part I : Distributed estimation of deterministic signals. *IEEE Trans. on Signal Proc.*, 56(1) :350–364, January 2008.
- [2] R. Olfati-Saber. Distributed kalman filtering for sensor networks. In *Proc. of the 46th IEEE Conf. on Decision and Control*, New Orleans, LA, USA, December 12–14 2007.
- [3] A. Fax and R. Murray. Information flow and cooperative control of vehicle formations. *IEEE Trans. on Automatic Control*, 49 :1465–1476, 2004.

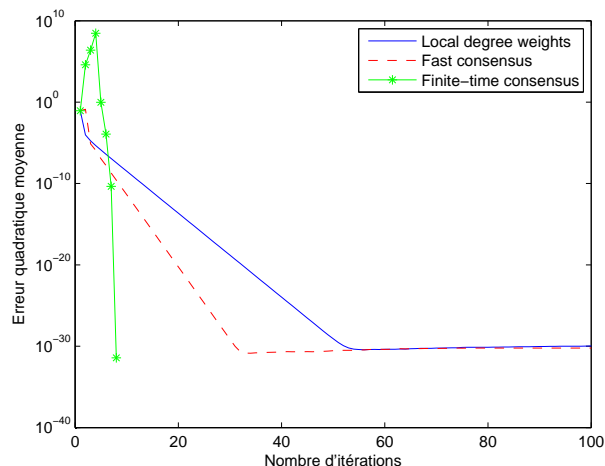


FIG. 2 – Comparaison de l'algorithme proposé avec les algorithmes standard et accéléré.

- [4] Lara Briñon Arranz, Alexandre Seuret, and Carlos Canudas De Wit. General Framework using Affine Transformations to Formation Control Design. In *Proceedings of 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys'10)*, Annecy France, Sept. 2010.
- [5] N. Lynch. *Distributed algorithms*. Morgan Kaufmann Publishers, San Mateo, CA, 1996.
- [6] A.Y. Kibangu and A.L.F. de Almeida. Distributed PARAFAC based DS-CDMA blind receiver for wireless sensor networks. In *Proc. of the IEEE Workshop SPAWC*, Marrakech, Morocco, June 20-23 2010.
- [7] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009.
- [8] F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. In Alberto Bemporad, Maurice Heemels, and Mikael Johansson, editors, *Networked Control Systems*, volume 406 of *Lecture Notes in Control and Information Sciences*, pages 75–107. Springer, 2011.
- [9] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems Control Lett.*, 53 :65–78, 2004.
- [10] E. Kokiopoulou and P. Frossard. Accelerating distributed consensus using extrapolation. *IEEE Signal Proc. Letters*, 14(10) :665–668, October 2007.
- [11] D.B. Kingston and R.W. Beard. Discrete-time average consensus under switching network topologies. In *Proc. of American Control Conference (ACC)*, Minneapolis, Minnesota, USA, 2006.
- [12] S. Sundaram and C. Hadjicostis. Distributed function calculation and consensus using linear iterative strategies. *IEEE Journal on Selected Areas in Communications*, 26(4) :650–660, May 2008.
- [13] S. Sundaram and C. Hadjicostis. Finite-time distributed consensus in graphs with time-invariant topologies. In *Proc. of American Control Conference (ACC)*, New York City, USA, July 11-13 2007.
- [14] J.-C. Delvenne, R. Carli, and S. Zampieri. Optimal strategies in the average consensus problem. *Systems and Control Letters*, 58 :759–765, 2009.
- [15] N. Maréchal. *Consensus de moyenne dans les réseaux de capteurs : Applications et Optimisation de performances*. PhD thesis, Institut National des Sciences Appliquées de Lyon, Lyon, France, 2009.
- [16] C.-K. Ko. *On matrix factorization and scheduling for finite-time average consensus*. PhD thesis, California Institute of Technology, Pasadena, California, USA, January 2010.
- [17] A.Y. Kibangu. Finite-time average consensus based protocol for distributed estimation over awgn channels. Submitted to CDC, 2011.