Sparse Solutions to the Average Consensus Problem via Various Regularizations of the Fastest Mixing Markov-Chain Problem

Giorgio Gnecco, Rita Morisi, and Alberto Bemporad, *Fellow, IEEE*

Abstract—In the consensus problem on multi-agent systems, in which the states of the agents represent opinions, the agents aim at reaching a common opinion (or consensus state) through local exchange of information. An important design problem is to choose the degree of interconnection of the subsystems to achieve a good trade-off between a small number of interconnections and a fast convergence to the consensus state, which is the average of the initial opinions under mild conditions. This paper addresses this problem through l_1-norm and l_0-"pseudo-norm" regularized versions of the well-known Fastest Mixing Markov-Chain (FMMC) problem. We show that such versions can be interpreted as robust forms of the FMMC problem and provide results to guide the choice of the regularization parameter.

Index Terms—Consensus, sparsity, optimization, regularization, Fastest Mixing Markov-Chain problem

1 Introduction

Several complex dynamical systems (e.g., wireless sensor networks, robotic teams, social networks) can be decomposed into a large number of subsystems (or agents), whose interactions are local and can be modeled by weighted edges in a communication graph, in which the vertices are the subsystems. Control problems on such multi-agent systems enjoy properties related to the structure of the communication graph, described, e.g., by weighted/unweighted adjacency and graph-Laplacian matrices [1], [2]. A paradigmatic example of such control problems is the consensus problem [3], in which the states of the subsystems represent opinions, and the agents aim at reaching a common opinion (or consensus state) through local exchange of information, without any form of centralization. A typical example is distributed estimation in wireless sensor networks [4]. Under mild conditions, one can prove that the consensus state is the average of the initial opinions, and the problem is called the average consensus problem [3]. In both problems, the variables to be chosen are the weights to be assigned to the edges of the communication graph. Such weights define a weighted adjacency matrix and, in the case of undirected communication graphs, also a weighted Laplacian matrix, whose spectral properties (i.e., properties expressed in terms of the eigenvalues/eigenvectors of such matrices) determine the rate of convergence to the consensus state [3], [5]. Interestingly, in the undirected case, determining the weights that optimize such spectral properties can be formulated as a convex optimization problem [5] (specifically, as a semidefinite program (SDP)), which is known as the Fastest Mixing Markov-Chain problem (Problem FMMC, in the paper) and can be solved efficiently, e.g., using interior-point methods. In several cases, one may be interested to find a suitable compromise between the desired spectral properties of the graph and the number of non-zero weights of the edges, thus obtaining sparse solutions to the consensus/average consensus problems with satisfactory rate of convergence. This is motivated, e.g., in the case of a high cost of communication associated with each edge. In this context, some recent results towards such sparse solutions are provided in [6], which investigates sparse graphs with certain symmetries, for which one has closed-form expressions for the Laplacian eigenvalues. However, the results are restricted to specific kinds of graphs. Another recent result in this direction is provided by [7], which shows that every Laplacian matrix associated with a symmetric graph can be well-approximated by the Laplacian matrix of a sparse subgraph, thus keeping the desired spectral properties but with a small number of edges. In principle, such a property could be exploited to find a good sparse solution to the average consensus problem, starting from a dense graph with good spectral properties in terms of the rate of convergence to the consensus state, then sparsifying such a graph, trying to preserve such spectral properties.

The present work focuses on the average consensus problem and, in the case of undirected graphs, proposes a different trade-off between good spectral properties of the communication graph and its sparsity, using an approach based on an l_1-norm regularized version of Problem FMMC, which is called Problem FMMC-$l_1(\eta)$ in the paper, where $\eta > 0$ denotes the regularization parameter. This variation of Problem FMMC is motivated by the fact that, due to geometrical properties of the l_1-norm [8], the introduction of such a regularization term in the objective of a convex optimization problem often enforces the sparsity of an optimal solution of the regularized version of that problem. We also consider another variation of Problem

- *The authors are with the IMT - Institute for Advanced Studies, Piazza S. Francesco 1955100, Lucca, Italy.*
E-mail: {giorgio.gnecco, rita.morisi, alberto.bemporad}@imtlucca.it.

Manuscript received 6 May 2015; revised 28 Aug. 2015; accepted 1 Sept. 2015. Date of publication 14 Sept. 2015; date of current version 13 Oct. 2015. Recommended for acceptance by M. Franceschetti.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TNSE.2015.2479086

FMMC (called Problem $\text{FMMC}_{\text{constr}}$-$l_1(\eta)$ in the paper) in which, besides the introduction of the l_1-norm regularization term, the weights of some edges are fixed, and another one in which the l_1-norm regularization term is replaced by a l_0-pseudo-norm regularization term. Then, we provide both a theoretical analysis of such variations of Problem FMMC and a numerical example modeling a wireless sensor network, comparing their solutions with the one obtained solving Problem FMMC. Related approaches were proposed in [9, Section 7.2] and in [10]. In particular, Problems FMMC-$l_1(\eta)$ and $\text{FMMC}_{\text{constr}}$-$l_1(\eta)$ are similar to one already proposed and investigated numerically in [9, Section 7.2], with the difference that the l_1-norm term in that reference appears inside an additional constraint instead than in the objective. We also mention that, for the average consensus problem in the presence of disturbances, a similar graph-sparsification optimization problem was also recently considered in [10], and solved through the Alternating Direction Method of Multipliers (ADMM) [11]. However, up to our knowledge, the theoretical analysis of Problems FMMC-$l_1(\eta)$ and $\text{FMMC}_{\text{constr}}$-$l_1(\eta)$ presented in this paper includes novel theoretical contributions in Sections 4c), d), f), g). In more details, Sections 4c), d), f), g) contain theoretical results that are specific to the l_1-regularized Problem FMMC and were not derived in [9]. In particular, to the best of our knowledge, our interpretation of Problem FMMC-$l_1(\eta)$ as a robust version of the Fastest Mixing Markov-Chain problem is novel, together with the other theoretical results we have obtained in Section 4 using Gershgorin's theorem and Weyl's inequalities. Instead, Sections 4 a), b) and e) provide results common also to l_1-norm regularizations of other convex optimization problems (and reported in the paper for completeness, and for their applicability to Problems FMMC-$l_1(\eta)$ and $\text{FMMC}_{\text{constr}}$-$l_1(\eta)$), whereas Section 4$h$) provides semidefinite programming formulations similar to the one presented in [9, Section 7.2], which are useful for solving Problems FMMC-$l_1(\eta)$ and $\text{FMMC}_{\text{constr}}$-$l_1(\eta)$ numerically. Finally, in Section 5, we also investigate theoretically another regularized version of Problem FMMC, called Problem FMMC-$l_0(\eta)$, whose analysis presented in that section is novel.

The paper extends significantly its conference version [12], including the proofs of all the theoretical results already stated therein (Propositions 2 and 3 of this work), and including and proving new theoretical results (Propositions 1, 4, and all the results in Section 5 of this work). Also the section about the numerical results has been extended significantly.

The paper is organized as follows. Section 2 summarizes the FMMC problem and introduces its equivalent formulation. Then, Section 3 presents two modifications of such a formulation (both obtained adding an l_1-norm regularization term to enforce sparsity, and fixing also some weights in the second one), which are investigated in Section 4 from a theoretical point of view. Section 5 extends the analysis to an l_0-pseudo-norm regularized version of the FMMC problem, for which interesting properties hold, despite the combinatorial nature of such a regularized optimization problem. Section 6 shows the application of the theoretical results obtained in the paper to the design of a wireless sensor network, and compares the regularized versions of the FMMC problem numerically. Finally, Section 7 discusses possible extensions of the work.

2 The Fastest Mixing Markov-Chain Problem

The consensus problem consists in determining the strengths of the interconnections among the subsystems of a multi-agent system, so that their states converge to a common state, subject to given topological constraints on the admissible connections. In the simplest case, the subsystems are linear, their states $x_i \in \mathbb{R}$ are scalar-valued, and the evolution of each subsystem i is determined by the discrete-time dynamics

$$x_i(t+1) = \sum_{j=1}^{n} P_{ij} x_j(t)\,, t = 0, 1, \dots\,, \tag{1}$$

where $P \in \mathbb{R}^{n \times n}$ is a matrix of interconnections with non-negative entries, satisfying the conditions $P\underline{1}_n = \underline{1}_n$ (here, $\underline{1}_n \in \mathbb{R}^n$ denotes a column vector of dimension n whose components are all equal to 1) and

$$P_{ij} = 0, \text{ if } i \neq j \text{ and } (i,j) \notin \mathcal{E}\,, \tag{2}$$

where $\mathcal{E}$ is a given set of admissible interconnections. In a design phase, the elements of the matrix P can be chosen arbitrarily, provided that the conditions above on P are satisfied.

The non-negativity assumption on P, together with the condition $P\underline{1}_n = \underline{1}_n$, implies that 1 is the eigenvalue of the matrix P with maximum absolute value (this can be proved, e.g., by an application of Gersghorin's theorem)[1], and that the state $x_i(t+1)$ at time $t+1$ is a convex combination of the states $x_j(t)$ at time t.

It is well known (see, e.g., [3]) that, when the eigenvalue 1 has algebraic multiplicity equal to 1, and all the other eigenvalues of P have absolute value smaller than 1, the states of the subsystems converge to the same consensus state x_c, when $t \to \infty$:

$$\lim_{t\to\infty} x_i(t) = \sum_{j=1}^{n} \alpha_j x_j(0) := x_c\,, \text{ for all } i \in \{1, \dots, n\}\,, \tag{3}$$

where α_j, $j = 1, \dots, n$, are suitable non-negative constants such that $\sum_{j=1}^{n} \alpha_j = 1$. When P is symmetric, one can show (see, e.g., [3]) that

$$\alpha_j = \frac{1}{n}\,, \forall j \in \{1, \dots, n\}\,, \tag{4}$$

and the consensus state is simply the average of the initial states (in such case, the problem is called the average consensus problem). In the following, we will focus on such a situation, therefore assuming $P = P^T$.

A particularly important aspect of the average consensus problem is the rate of convergence to the average consensus state, which is related to the second-largest eigenvalue modulus of P:

$$\mu(P) := \max_{j=2,\dots,n} |\lambda_j\{P\}|\,, \tag{5}$$

1. Gersghorin's theorem [13, Section 7.2] states that all the eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$ belong, in the complex plane, to at least one of the Gersghorin circles G_i (for $i = 1, \dots, n$), whose centers and radii are defined, respectively, by A_{ii} and $\sum_{j \neq i,\, j=1}^{n} |A_{ij}|$.

where the (real) eigenvalues $\lambda_j\{P\}$, $j = 1, \ldots, n$, have been ordered with their multiplicity in a nonincreasing order, i.e., $1 = \lambda_1\{P\} \geq \lambda_2\{P\} \geq \cdots \lambda_j(P) \geq \cdots \geq \lambda_n\{P\} > -1$. The smaller $\mu(P)$, the faster the convergence to the consensus state [3].

In addition, a related quantity is the mixing time [5]

$$\tau(P) := \frac{1}{\log\left(\frac{1}{\mu(P)}\right)}\,, \tag{6}$$

which is an asymptotic measure of the number of steps required for reducing by the Euler's number e a suitable distance (the total variation distance) between the global state vector and the vector whose components are equal to the average consensus state.

Since the symmetric matrix P has non-negative elements and satisfies $P\underline{1}_n = \underline{1}_n$, its generic element P_{ij} can be interpreted as a transition probability from the vertex i to the vertex j of a graph (including the case of a self-loop when $i = j$), whose vertices are the subsystems. Hence, the rate of convergence of the Markov chain with transition probabilities P_{ij} to its stationary distribution depends on $\mu(P)$.

The problem of determining the coefficients P_{ij} that minimize $\mu(P)$ subject to a given topology of the graph is called the Fastest Mixing Markov-Chain problem (Problem FMMC, in the following), formulated as [5]

Problem FMMC (first formulation):

$$\begin{aligned} \underset{P \in \mathbb{R}^{n\times n}}{\text{minimize}} \quad & \mu(P) \\ \text{subject to} \quad & P\underline{1}_n = \underline{1}_n,\ P = P^T\,, \\ & P_{ij} \geq 0,\ \forall i, j \in \{1, \ldots, n\}\,, \\ & P_{ij} = 0,\ \text{if } (i,j) \notin \mathcal{E}\,. \end{aligned} \tag{7}$$

Interestingly, this is a convex optimization problem, since

$$\mu(P) = |\lambda|_{\max}\left\{P - \frac{1}{n}\underline{1}_n\underline{1}_n^T\right\} \tag{8}$$

(see [5] for a proof of formula (8)), where $|\lambda|_{\max}$ stands for the largest eigenvalue modulus. Moreover, Problem FMMC can also be written as a semidefinite program [5, Section 2.3].

We introduce an equivalent version of Problem FMMC, using a notation suitable for its sparse extensions presented in Section 3 and for their theoretical investigations in Sections 4 and 5.

In the following, we denote by $\underline{w} \in \mathbb{R}^m$ the column vector of weights associated with the m edges joining different vertices, and by $\underline{w}_{\text{sl}} \in \mathbb{R}^n$ the column vector of weights associated with the n self-loop edges. Hence, we can represent the weighted adjacency matrix P as a linear function $P(\underline{w}, \underline{w}_{\text{sl}})$ of such weights. For instance, for $n = 3$ and $m = n(n-1)/2$ (the case of a complete graph), one obtains the symmetric matrix

$$P(\underline{w}, \underline{w}_{\text{sl}}) = \begin{bmatrix} w_{\text{sl},1} & w_1 & w_2 \\ w_1 & w_{\text{sl},2} & w_3 \\ w_2 & w_3 & w_{\text{sl},3} \end{bmatrix}. \tag{9}$$

Moreover, introducing the vertex-edge incidence matrix $M \in \mathbb{R}^{n\times m}$, whose elements are defined as follows:

$$M_{ij} = \begin{cases} 1\,, & \text{if the vertex } i \text{ is an endpoint of} \\ & \text{the (non self-loop) edge } j\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{10}$$

and setting

$$\underline{w}_{\text{sl}} := \underline{1}_n - M\underline{w}\,, \tag{11}$$

the constraints

$$P_{ij} \geq 0 \text{ for any } i, j \in \{1, \ldots, n\} \text{ and } P\underline{1}_n = \underline{1}_n \tag{12}$$

are equivalent to

$$w_i \geq 0 \text{ for any } i \in \{1, \ldots, m\} \text{ and } M\underline{w} \leq \underline{1}_n\,. \tag{13}$$

Using (11), the matrix P becomes an affine function $P(\underline{w})$ of the weight vector $\underline{w}$, and the second-largest eigenvalue modulus of P is expressed as a convex function—denoted by $\mu(\underline{w})$—of the weight vector $\underline{w}$, since convexity is preserved by affine mappings [14, Section 3.2]. With the notations just introduced, Problem FMMC can be compactly rewritten as

Problem FMMC (second formulation):

$$\begin{aligned} \underset{\underline{w} \in \mathbb{R}^m}{\text{minimize}} \quad & f(\underline{w}) := \mu(\underline{w}) \\ \text{subject to} \quad & \underline{w} \geq \underline{0}_m, \\ & M\underline{w} \leqslant \underline{1}_n\,. \end{aligned} \tag{14}$$

Remark 1. For each admissible weight vector $\underline{w}$, the weighted adjacency matrix $P(\underline{w})$ can be interpreted as a "walk matrix" [15] on the weighted graph described by $P(\underline{w})$, for which it is well-known that its maximum (and maximum modulus) eigenvalue $\lambda_{\max}\{P(\underline{w})\} = 1$ has multiplicity equal to the number of connected components of such a graph. So, in case of a disconnected graph, for every weight vector $\underline{w}$, $\lambda_{\max}\{P(\underline{w})\} = 1$ has multiplicity at least 2, and, trivially, for the second-largest eigenvalue modulus, one has $\mu(\underline{w}) = 1$.

In order to avoid the pathological situation described in Remark 1, in the remaining of the paper we generally assume the graph described by the vertex-edge incidence matrix M to be connected.

3 Some Sparse Variations of the Fastest Mixing Markov-Chain problem

We now consider the following sparse variations of Problem FMMC.

3.1 Problem FMMC with a Regularization Term

In order to find a good compromise between sparsity of $\underline{w}$ and a small value of the second-largest eigenvalue modulus of the weighted adjacency matrix $P(\underline{w})$, we consider, for any $\eta > 0$, the following regularized version of Problem FMMC, in which an l_1-regularization term with regularization parameter η is added to the objective (here, $\|\underline{w}\|_1 := \sum_{i=1}^m |w_i|$):

Problem FMMC-$l_1(\eta)$:

$$\begin{aligned} \underset{\underline{w} \in \mathbb{R}^m}{\text{minimize}} \quad & f^{(1,\eta)}(\underline{w}) := \mu(\underline{w}) + \eta\|\underline{w}\|_1 \\ \text{subject to} \quad & \underline{w} \geq \underline{0}_m, \\ & M\underline{w} \leqslant \underline{1}_n\,. \end{aligned} \tag{15}$$

The term $\eta\|\underline{w}\|_1$ in (15) often induces sparsity of a resulting optimal solution $\underline{w}^\circ(\eta)$ [8], i.e., many components of $\underline{w}^\circ(\eta)$ tend to be 0.

In Section 5, we investigate also the regularized version of Problem FMMC obtained by replacing the l_1-norm in Problem FMMC-$l_1(\eta)$ with the l_0-"pseudo-norm"

$$\|\underline{w}\|_0 := \text{number of non-zero components of } \underline{w}\,. \tag{16}$$

We state such a problem as follows:

Problem FMMC-$l_0(\eta)$:

$$\begin{aligned} &\underset{\underline{w}\in\mathbb{R}^m}{\text{minimize}} && f^{(0,\eta)}(\underline{w}) := \mu(\underline{w}) + \eta\|\underline{w}\|_0 \\ &\text{subject to} && \underline{w} \geq \underline{0}_m, \\ & && M\underline{w} \leqslant \underline{1}_n\,. \end{aligned} \tag{17}$$

Although the l_0-pseudo-norm is a more natural way to enforce sparsity than the l_1-norm, it is a nonconvex function, so, when sparsity is desired, it is common to replace the l_0-pseudo-norm with the l_1-norm, which is a convex function.

3.2 Problem FMMC with Fixed Edges and a Regularization Term

An interesting variation of Problem FMMC-$l_1(\eta)$ consists in fixing some components of the weight vector $\underline{w}$. This is motivated, e.g., when one is interested in imposing some additional structure on the topology of the graph resulting from the optimization of the weight vector (e.g., enforcing the presence of given subgraphs, such as trees connecting important "backbone" vertices). Without loss of generality, in the following we assume (up to a permutation of the indices) that the fixed weights are the first m_{fixed} ones (where $1 \leq m_{\text{fixed}} \leq m$), whereas the last $m_{\text{free}} : m - m_{\text{fixed}}$ weights are not fixed (the special case $m_{\text{fixed}} = m$ is trivial). We then decompose the column vector $\underline{w}$ as

$$\underline{w} = \text{col}(\underline{w}_{\text{fixed}}, \underline{w}_{\text{free}}) \tag{18}$$

and the vertex-edge incidence matrix M as

$$M = [M_{\text{fixed}}|M_{\text{free}}]\,, \tag{19}$$

and we express the second-largest eigenvalue modulus μ as a function $\mu(\underline{w}_{\text{free}})$ of the unfixed weights only. Then, for a given choice of the weight vector $\underline{w}_{\text{fixed}}$, we consider the following optimization problem:

Problem FMMC$_{\text{constr}}$-$l_1(\eta)$:

$$\begin{aligned} &\underset{\underline{w}_{\text{free}}\in\mathbb{R}^{m_{\text{free}}}}{\text{minimize}} && f^{(1,\eta)}_{\text{constr}}(\underline{w}_{\text{free}}) := \mu(\underline{w}_{\text{free}}) + \eta\|\underline{w}_{\text{free}}\|_1 \\ &\text{subjectto} && \underline{w}_{\text{free}} \geq \underline{0}_{m_{\text{free}}}, \\ & && M_{\text{free}}\underline{w}_{\text{free}} \leqslant \underline{1}_n - M_{\text{fixed}}\underline{w}_{\text{fixed}}\,. \end{aligned} \tag{20}$$

Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ has a form which is similar to the one of Problem FMMC-$l_1(\eta)$. We assume in the following that the fixed weights have been chosen in such a way that the polyhedron

$$\{\underline{w} \in \mathbb{R}^{m_{\text{free}}} : \underline{w}_{\text{free}} \geq \underline{0}_{\text{free}}, M_{\text{free}}\underline{w}_{\text{free}} \leqslant \underline{1}_n - M_{\text{fixed}}\underline{w}_{\text{fixed}}\} \tag{21}$$

is non-empty, so that Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ admits a feasible solution.

Remark 2. A similar variation can be studied replacing the l_1-norm with the l_0-pseudo-norm, but it is not investigated here, to avoid redundancy in the analysis. Another variation is obtained assuming that some edges are just "sufficiently used" rather than "fixed", i.e., that, for some indices i and some constants $\beta_i \in [0,1]$, one has $w_i \geq \beta_i$. The resulting problem has still linear equality and inequality constraints.

Remark 3. Although the l_1-norm is nondifferentiable at the origin, the terms $\|\underline{w}\|_1$ and $\|\underline{w}_{\text{free}}\|_1$ in the objectives of Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$ can also be written, respectively, as $\underline{1}_m^T\underline{w}$ and $\underline{1}_{m_{\text{free}}}^T\underline{w}_{\text{free}}$ (thus, as linear—hence differentiable—terms), due to the respective non-negativity constraints $\underline{w} \geq \underline{0}_m$ and $\underline{w}_{\text{free}} \geq \underline{0}_{m_{\text{free}}}$.

In the next section we provide some theoretical results about the optimal solutions of Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$.

4 Theoretical Results for Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$

We first consider the analysis of Problem FMMC-$l_1(\eta)$; extensions of the results to Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ are considered later in this section.

a) Existence of an optimal solution.

The next result states the existence of an optimal solution to Problem FMMC-$l_1(\eta)$.

Proposition 1. *Problem FMMC-$l_1(\eta)$ admits an optimal solution for every $\eta > 0$.*

Proof. The feasible set of Problem FMMC-$l_1(\eta)$ is convex, closed, and bounded. Moreover, its objective is continuous since the l_1-norm regularization term $\|\underline{w}\|_1$ is continuous, and on the feasible set the second-largest eigenvalue modulus $\mu(P)$ has the expression (8), which is continuous due to the continuous dependence of the eigenvalues of a matrix on its entries [16, Section 7.6], and the fact that the point-wise maximum of a finite set of continuous functions is continuous, too. Concluding, Problem FMMC-$l_1(\eta)$ involves the minimization of a continuous objective function on a compact set, so an optimal solution to Problem FMMC-$l_1(\eta)$ exists by Weierstrass theorem. □

b) Effect of the regularization parameter.

Solving Problem FMMC-$l_1(\eta)$ involves finding a good compromise between the minimization of the term $\mu(\underline{w})$ and the one of $\|\underline{w}\|_1$. Next Proposition 2 shows that the regularization parameter η has opposite effects on the two terms $\mu(\underline{w})$ and $\|\underline{w}\|_1$, when evaluated at an optimal solution.

Proposition 2. *Let $0 < \eta_1 < \eta_2$, and $\underline{w}_1^\circ(\eta_1), \underline{w}_1^\circ(\eta_2)$ be optimal solutions to Problem FMMC-$l_1(\eta_1)$ and Problem FMMC-$l_1(\eta_2)$, respectively. Then,*

i) $\mu(\underline{w}_1^\circ(\eta_1)) \leq \mu(\underline{w}_1^\circ(\eta_2))$,

ii) $\|\underline{w}_1^\circ(\eta_1)\|_1 \geq \|\underline{w}_1^\circ(\eta_2)\|_1$.

Proof. By the optimality of $\underline{w}_1^\circ(\eta_1)$ for Problem FMMC-$l_1(\eta_1)$, one has

$$\mu(\underline{w}_1^\circ(\eta_1)) + \eta_1\|\underline{w}_1^\circ(\eta_1)\|_1 \leq \mu(\underline{w}_1^\circ(\eta_2)) + \eta_1\|\underline{w}_1^\circ(\eta_2)\|_1\,. \tag{22}$$

Similarly, by the optimality of $\underline{w}_1^\circ(\eta_2)$ for Problem FMMC-$l_1(\eta_2)$, one gets

$$\mu(\underline{w}_1^\circ(\eta_2)) + \eta_2\|\underline{w}_1^\circ(\eta_2)\|_1 \leq \mu(\underline{w}_1^\circ(\eta_1)) + \eta_2\|\underline{w}_1^\circ(\eta_1)\|_1 . \quad (23)$$

Combining the two inequalities above, one obtains

$$\begin{aligned}&\eta_2(\|\underline{w}_1^\circ(\eta_2)\|_1 - \|\underline{w}_1^\circ(\eta_1)\|_1)\\ &\leq \mu(\underline{w}_1^\circ(\eta_1)) - \mu(\underline{w}_1^\circ(\eta_2))\\ &\leq \eta_1(\|\underline{w}_1^\circ(\eta_2)\|_1 - \|\underline{w}_1^\circ(\eta_1)\|_1),\end{aligned}$$

which is satisfied if and only if conditions i) and ii) hold, as $0 < \eta_1 < \eta_2$. □

In general, instead, the sparsity

$$\begin{aligned}s(\underline{w}_1^\circ(\eta)) &:= 1 - \|\underline{w}_1^\circ(\eta)\|_0/m\\ &= \text{ fraction of zero components of } \underline{w}^\circ(\eta)\end{aligned}$$

of an optimal solution to Problem FMMC-$l_1(\eta)$ may not be a monotonic function of η, as shown in Fig. 5 (see Section 6). This behavior is similar to the one observed for other l_1-regularized optimization problems, such as the classical Least Absolute Shrinkage and Selection Operator (LASSO) problem (see., e.g., [17, Fig. 1]).

c) Conditions under which $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_1(\eta)$.

The next result states conditions on the regularization parameter under which $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_1(\eta)$, or its unique optimal solution. An application of the result to the choice of the regularization parameter is given in Section 4*e*).

Proposition 3. *Let $\eta \geq 2$. Then $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_1(\eta)$. If $\eta > 2$, then $\underline{w} = \underline{0}_m$ is its unique optimal solution.*

Proof. Let $\Delta\underline{w}$ be an arbitrary admissible increment of $\underline{w}$, starting from $\underline{w} = \underline{0}_m$ (such an increment can be used to generate the whole set of admissible solutions to Problem FMMC-$l_1(\eta)$, since it is convex and contains $\underline{0}_m$). Then, the corresponding increment $\Delta f^{(1,\eta)}$ in the objective $f^{(1,\eta)}$ of Problem FMMC-$l_1(\eta)$ is

$$\Delta f^{(1,\eta)} = \mu(\Delta\underline{w}) - \mu(\underline{0}_m) + \eta\|\Delta\underline{w}\|_1 . \quad (24)$$

Now, one has $\mu(\underline{0}_m) = 1$ (as the associated weighted adjacency matrix is $P(\underline{0}_m) = I_n$, the identity matrix of dimension $n \times n$), whereas one can find a lower bound on $\mu(\Delta\underline{w})$ as follows. The matrix $P(\Delta\underline{w}))$ can be written as

$$P(\Delta\underline{w}) = I_n + E(\Delta\underline{w}) , \quad (25)$$

where the main-diagonal entries of the matrix $E(\Delta\underline{w})$ are non-positive with their absolute values bounded from above by $\|\Delta\underline{w}\|_1$, whereas, for each row i, one has

$$\sum_{j\neq i,\, j=1}^{n} |E_{ij}(\Delta\underline{w})| \leq \|\Delta\underline{w}\|_1 . \quad (26)$$

Then, by Gersghorin's theorem (recall footnote 1), all the eigenvalues of $E(\Delta\underline{w})$ are bounded from above in absolute value by $2\|\Delta\underline{w}\|_1$. As the presence of the matrix I_n in formula (25) has only the effect of translating the eigenvalues of $E(\Delta\underline{w})$ by 1, one finally obtains

$$\mu(\Delta\underline{w}) \geq 1 - 2\|\Delta\underline{w}\|_1 . \quad (27)$$

Concluding, for an arbitrary admissible increment $\Delta\underline{w}$, one gets

$$\Delta f^{(1,\eta)} \geq 1 - 2\|\Delta\underline{w}\|_1 - 1 + \eta\|\Delta\underline{w}\|_1 = -2\|\Delta\underline{w}\|_1 + \eta\|\Delta\underline{w}\|_1 . \quad (28)$$

If $\eta \geq 2$, then $\Delta f^{(1,\eta)}$ is non-negative for every admissible $\Delta\underline{w}$, hence $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_1(\eta)$. If $\eta > 2$, then $\Delta f^{(1,\eta)}$ is positive for every $\Delta\underline{w} \neq \underline{0}_m$, hence $\underline{w} = \underline{0}_m$ is the only optimal solution to Problem FMMC-$l_1(\eta)$. □

The following example shows that the bound obtained in Proposition 3 is tight, at least if one does not impose further restrictions on the class of graphs to be considered.

Example 1. Let $n = 2$ and $m = 1$. Then, the matrix $P(\underline{w})$ has the expression

$$P(\underline{w}) = \begin{bmatrix} 1 - w_1 & w_1 \\ w_1 & 1 - w_1 \end{bmatrix}, \quad (29)$$

whose eigenvalues are 1 and $1 - 2w_1$. Hence, on the set $[0, 1]$ of admissible solutions to Problem FMMC-$l_1(\eta)$, its objective is $|1 - 2w_1| + \eta w_1$, and $w_1 = 0$ is, respectively, the unique optimal solution to Problem FMMC-$l_1(\eta)$ for $\eta > 2$, one of its (infinite) optimal solutions for $\eta = 2$ (the optimal ones being all $w_1 \in [0, \frac{1}{2}]$), and a suboptimal solution for $0 < \eta < 2$ (the optimal one being $w_1 = \frac{1}{2}$).

d) Non differentiability of the objective at $\underline{w} = \underline{0}_m$.

In the proof of Proposition 3, we have used Gersghorin's theorem instead than an approach based on first-order optimality conditions for a differentiable objective. The reason is that, as shown in the next Proposition 4, in general the objective of Problem FMMC-$l_1(\eta)$ is not differentiable at $\underline{w} = \underline{0}_m$, although the l_1-regularization term is represented by a linear function on its set of admissible solutions. To state Proposition 4, we need the following definition.

Definition 1. *A function $y : D \to \mathbb{R}$ defined on a convex and compact subset $D \subseteq \mathbb{R}^m$ is differentiable at a point $\underline{x}_0$ belonging to the boundary ∂D of D iff there exists a linear map $J : \mathbb{R}^m \to \mathbb{R}$ such that*

$$\lim_{t\to 0^+} \frac{y(\underline{x}_0 + t\underline{h}) - y(\underline{x}_0) - J(t\underline{h})}{\|t\underline{h}\|} = 0 , \quad (30)$$

for all feasible directions $\underline{h} \in \mathbb{R}^m$, i.e., such that $\underline{x}_0 + t\underline{h} \in D$ for all $t > 0$ sufficiently small.

Remark 4. In the usual definition of differentiability, instead, one assumes that the point $\underline{x}_0$ belongs to the interior of the set D, and in that case all vectors $\underline{h} \in \mathbb{R}^m$ are feasible directions. In Definition 1, we extend this concept to a point belonging to the boundary of the domain (which is the case of $\underline{w} = \underline{0}_m$ for Problem FMMC-$l_1(\eta)$, as it belongs to the boundary of the domain of the function $\mu(\underline{w})$).

Remark 5. For $i = 1, \ldots, n$, let $\underline{e}_i \in \mathbb{R}^m$ denote the vector whose component i is equal to 1, and all its other components are equal to 0. If the directions $\underline{e}_i$ are feasible, (30) and the linearity of the map J imply, for every feasible direction $\underline{h}$,

$$\lim_{t \to 0^+} \frac{y(\underline{x}_0 + t\underline{h}) - y(\underline{x}_0)}{\|t\underline{h}\|} = \sum_{i=1}^{m} h_i \lim_{t \to 0^+} \frac{y(\underline{x}_0 + t\underline{e}_i) - y(\underline{x}_0)}{t}.$$

Proposition 4. *Let $n \geq 3$. Then the objective of Problem FMMC-$l_1(\eta)$ is nondifferentiable at $\underline{w} = \underline{0}_m$.*

Proof. For $i = 1, \ldots, n$ and for each $t \in [0, 1]$, the graph associated with the weighted adjacency matrix $P(t\underline{e}_i)$ is disconnected, hence the maximum (and maximum modulus) eigenvalue $\lambda_{\max}\{P(t\underline{e}_i)\} = 1$ of $P(t\underline{e}_i)$ has multiplicity at least 2, and its second-largest eigenvalue modulus is $\mu(t\underline{e}_i) = 1$. Thus, at $\underline{w} = \underline{0}_m$, the directional derivative of the objective of Problem FMMC-$l_1(\eta)$ in the direction $\underline{e}_i$, i.e.,

$$\lim_{t \to 0^+} \frac{\mu(t\underline{e}_i) - \mu(\underline{0}_m) + \eta \underline{1}_m^T (t\underline{e}_i)}{t}, \tag{31}$$

is equal to η. Now, let $\hat{\underline{w}}$ be any admissible weight vector with $\hat{\underline{w}} > \underline{0}_m$ (elementwise). Since the graph associated with the weighted adjacency matrix $P(\hat{\underline{w}})$ is connected, $\lambda_{\max}\{P(\hat{\underline{w}})\} = 1$ has multiplicity equal to 1, hence $\lambda_2(P(\hat{\underline{w}})) \leq \mu(\hat{\underline{w}}) < 1$, and $\lambda_n\{P(\hat{\underline{w}})\} > -1$ by Gersghorin's theorem. Now, for any $t \in [0, 1]$, one has $P(t\hat{\underline{w}}) = (1 - t)I + tP(\hat{\underline{w}})$. Then, the eigenvalues of $P(\hat{\underline{w}})$ and $P(t\hat{\underline{w}})$, ordered nonincreasingly with their multiplicity, are related by

$$\lambda_i\{P(t\hat{\underline{w}})\} = (1 - t) + t\lambda_i\{P(\hat{\underline{w}})\}, \quad (i = 1, \ldots, n), \tag{32}$$

so, assuming without loss of generality $\mu(\hat{\underline{w}}) = \lambda_2(P(\hat{\underline{w}})) > 0$ (the case $\mu(\hat{\underline{w}}) = \max\{|\lambda_2(P(\hat{\underline{w}}))|, |\lambda_n(P(\hat{\underline{w}}))|\}$ is similar), for $t > 0$ sufficiently small one gets $\mu(t\hat{\underline{w}}) = 1 - t + t\lambda_2(P(\hat{\underline{w}}))$. Hence, at $\underline{w} = \underline{0}_m$, the directional derivative of the objective of Problem FMMC-$l_1(\eta)$ in the direction $\hat{\underline{w}}$, i.e.,

$$\lim_{t \to 0^+} \frac{\mu(t\hat{\underline{w}}) - \mu(\underline{0}_m) + \eta \underline{1}_m^T (t\hat{\underline{w}})}{t}, \tag{33}$$

is equal to $\lambda_2(P(\hat{\underline{w}})) - 1 + \eta\|\hat{\underline{w}}\|_1$, which differs from

$$\sum_{i=1}^{m} \hat{w}_i \left(\lim_{t \to 0^+} \frac{\mu(t\underline{e}_i) - \mu(\underline{0}_m) + \eta \underline{1}_m^T (t\underline{e}_i)}{t} \right) = \eta\|\hat{\underline{w}}\|_1 \tag{34}$$

(due to Definition 1 and Remark 5, formulas (33) and (34) would have coincided, instead, in the case of differentiability of the objective of Problem FMMC-$l_1(\eta)$ at $\underline{w} = \underline{0}_m$). Then, we conclude that the objective of Problem FMMC-$l_1(\eta)$ is nondifferentiable at $\underline{w} = \underline{0}_m$. □

Remark 6. For $n = 2$ and $m = 1$, instead, the objective of Problem FMMC-$l_1(\eta)$ is differentiable at $\underline{w} = \underline{0}_m$, as shown by Example 1. This is not in contrast with the proof of Proposition 4, since in this case - as being $n < 3$—the graph associated with the weighted adjacency matrix $P(t\underline{e}_1)$ is connected for every $t \in (0, 1]$.

e) Choice of the regularization parameter and reoptimization.
The theoretical results presented above justify the following practical rule for choosing the regularization parameter η:

- given a positive integer N and a maximal acceptable increase $\varepsilon > 0$ for the second-largest eigenvalue modulus of P with respect to its optimal value μ°_{FMMC} in Problem FMMC, solve Problem FMMC-$l_1(\eta)$ in correspondence of N values $\eta^{(j)}$ for η such that
  - $0 < \eta^{(1)} < \eta^{(2)} < \cdots < \eta^{(N)} < 2$ (the last inequality needed to avoid just the trivial optimal solution $\underline{w}^\circ = \underline{0}_m$), and
  - $\mu(\underline{w}^\circ_1(\eta^{(j)})) \leq \mu^\circ_{FMMC} + \varepsilon \quad (j = 1, \ldots, N)$ (the inequality needed just to guarantee the desired tolerance on $\mu(\underline{w}^\circ_1(\eta^{(j)}))$, hence on the rate of convergence to the consensus state);
- choose $j^\circ \in \{1, \ldots, N\}$ that maximizes the sparsity $s(\underline{w}^\circ_1(\eta^{(j)}))$;
- perform a final reoptimization step, solving Problem FMMC on the graph obtained removing from the original graph all the non self-loop edges i for which $w^\circ_{1,i}(\eta^{(j^\circ)}) = 0$, obtaining another weight vector $\underline{w}^\circ_{\text{reopt}}$.

By construction, the last reoptimization step satisfies $\mu(\underline{w}^\circ_{\text{reopt}}) \leq \mu(\underline{w}^\circ_1(\eta^{(j^\circ)}))$ (due to the optimality of $\underline{w}^\circ_{\text{reopt}}$ on Problem FMMC on the new graph, and the feasibility of $\underline{w}^\circ_1(\eta^{(j^\circ)})$ for such an optimization problem), and $s(\underline{w}^\circ_{\text{reopt}}) \geq s(\underline{w}^\circ_1(\eta^{(j^\circ)}))$. Such a reoptimization step is common to other l_1-regularized optimization problems: for the LASSO, it is known as debiasing [18, Section 13.3.5].

Finally, a possible way to choose the tolerance parameter ε (which has to be in any case smaller than $1 - \mu^\circ_{FMMC}$, again to avoid trivial optimal solutions) consists in expressing it in terms of the maximal allowable ratio ρ between the mixing time $\tau(P(\underline{w}))$ and its optimal value $\tau^\circ_{FMMC} := \frac{1}{\log\left(\frac{1}{\mu^\circ_{FMMC}}\right)}$, which is obtained when solving Problem FMMC, i.e., one sets

$$\varepsilon = \left(\mu^\circ_{FMMC}\right)^{\frac{1}{\rho}} - \mu^\circ_{FMMC}. \tag{35}$$

f) Interpretation of Problem FMMC-$l_1(\eta)$ as a robust version of Problem FMMC.
Problem FMMC-$l_1(\eta)$ has also the following interpretation. Let us suppose that, for any given "nominal" choice of the weights w_i $(i = 1, \ldots, m)$, one has an "uncertainty" Δw_i such that $|\Delta w_i| \leq \delta|w_i|$, for some fixed $\delta > 0$. Then, an application of Gersghorin's theorem and Weyl's inequalities[2] in matrix-perturbation theory shows that the second-largest eigenvalue modulus $\mu(\underline{w} + \Delta\underline{w})$ is bounded from above as

2. Let $A, B \in \mathbb{R}^{n \times n}$ and symmetric, and let their eigenvalues be ordered nonincreasingly with their multiplicity as

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_j(A) \geq \cdots \geq \lambda_n(A), \tag{36}$$

$$\lambda_1(B) \geq \lambda_2(B) \geq \cdots \geq \lambda_j(B) \geq \cdots \geq \lambda_n(B). \tag{37}$$

Then, in their simplest form, Weyl's inequalities [19, Theorem 8.4.11] state that, for every $j = 1, \ldots, n$, one has

$$\lambda_j(A) + \lambda_n(B) \leq \lambda_j(A + B) \leq \lambda_j(A) + \lambda_1(B). \tag{38}$$

$$\mu(\underline{w} + \Delta\underline{w}) \leq \mu(\underline{w}) + 2\delta\|\underline{w}\|_1 . \tag{39}$$

Then, an optimal "robust" choice of the nominal weight vector $\underline{w}$ is obtained minimizing the objective $\mu(\underline{w}) + 2\delta\|\underline{w}\|_1$ on the set of admissible weight vectors $\underline{w}$, i.e., solving a robust version of Problem FMMC which takes into account the uncertainty of the weights, and is obtained replacing its objective $\mu(\underline{w})$ with $\mu(\underline{w}) + 2\delta\|\underline{w}\|_1$. However, this is equivalent to solving Problem FMMC-$l_1((\eta)$ with the choice $\eta = 2\delta$. Finally, we notice that, when $\delta \geq 1$, for every admissible nominal choice of the vector $\underline{w}$, the perturbation $\Delta\underline{w} = -\underline{w}$ is admissible for the robust version of Problem FMMC just described, and the resulting perturbed vector is $\underline{w} + \Delta\underline{w} = \underline{0}_m$, which satisfies $\mu(\underline{0}_m) = 1$, as being the resulting graph disconnected. Hence, when $\delta \geq 1$, for every nominal choice of $\underline{w}$ one cannot have $\mu(\underline{w} + \Delta\underline{w}) < 1$ for every admissible perturbation, and $\underline{w} = \underline{0}_m$ is an optimal nominal choice. This is consistent with Proposition 3.

g) Extension to Problem FMMC$_{\text{constr}}$-$l_1(\eta)$.

Apart from the tightness of the bound on the minimal value of the regularization parameter η for which $\underline{w}_{\text{free}} = \underline{0}_{m_{\text{free}}}$ is an optimal solution, the results above can be extended to Problem FMMC$_{\text{constr}}$-$l_1(\eta)$. In particular, Propositions 1, 2, and 3 can be extended to Problem FMMC$_{\text{constr}}$-$l_1(\eta)$, simply replacing $\underline{w}_{\text{free}}$ with $\underline{w}$. For the first two propositions, the extension requires no significant changes in the proofs. In the third case, the only significant change in the proof is the additional use of the above-mentioned Weyl's inequalities to get a formula similar to (27), bounding the eigenvalues of the sum of two symmetric matrices. To obtain the extension of Proposition 4, maintaining the structure of the proof, one requires the additional assumption that the subgraph containing only the fixed edges is disconnected, and remains disconnected when adding arbitrarily only one of the "free" edges.

h) Formulation through semidefinite programming.

Likewise Problem FMMC, Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$ can be formulated as semidefinite programs, allowing the use of interior-point methods for finding their optimal solutions. More precisely, one obtains the following alternative formulation[3] of Problem FMMC-$l_1(\eta)$, expressing the m non self-loop edges in terms of their endpoints as (i,j), and considering the set

$$\begin{aligned}\mathcal{E} := \{(i,j) : i \neq j,\ i,j \in \{1,\ldots,n\},\ \text{and}\\ \exists k \in \{1,\ldots,n\} \text{ such that } M_{ik} = M_{jk} = 1\}.\end{aligned} \tag{40}$$

Problem FMMC-$l_1(\eta)$ (SDP formulation):

$$\begin{aligned}&\underset{s\in\mathbb{R},P\in\mathbb{R}^{n\times n}}{\text{minimize}} && \left(s + \tfrac{\eta}{2}\sum_{i\neq j,\, i,j=1}^{n} P_{ij}\right)\\ &\text{subject to} && -sI \preceq P - \tfrac{1}{n}\underline{1}_n\underline{1}_n^T \preceq sI,\\ &&& P\underline{1}_n = \underline{1}_n,\ P = P^T,\\ &&& P_{ij} \geq 0,\ \forall i,j \in \{1,\ldots,n\},\\ &&& P_{ij} = 0,\ \text{if } (i,j) \notin \mathcal{E}.\end{aligned} \tag{41}$$

3. Here and in the following, for any two symmetric matrices $X, Y \in \mathbb{R}^{n\times n}$, the notation $X \preceq Y$ (resp., $X \succeq Y$) means that $X - Y$ is negative (resp., positive) semidefinite, whereas $X \leq Y$ (resp., $X \geq Y$) stands for an elementwise inequality.

One gets also the following alternative formulation of Problem FMMC$_{\text{constr}}$-$l_1(\eta)$, introducing the subset $\mathcal{E}_{\text{fixed}} \subseteq \mathcal{E}$ of edges (i,j) associated with fixed weights $P_{ij,\text{fixed}} = P_{ji,\text{fixed}}$.

Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ (SDP formulation):

$$\begin{aligned}&\underset{s\in\mathbb{R},P\in\mathbb{R}^{n\times n}}{\text{minimize}} && \left(s + \tfrac{\eta}{2}\sum_{i\neq j,\, i,j=1}^{n} P_{ij}\right)\\ &\text{subjectto} && -sI \preceq P - \tfrac{1}{n}\underline{1}_n\underline{1}_n^T \preceq sI,\\ &&& P\underline{1}_n = \underline{1}_n,\ P = P^T,\\ &&& P_{ij} \geq 0,\ \forall i,j \in \{1,\ldots,n\},\\ &&& P_{ij} = 0,\ \text{if } (i,j) \notin \mathcal{E},\\ &&& P_{ij} = P_{ij,\text{fixed}},\ \text{if } (i,j) \in \mathcal{E}_{\text{fixed}}.\end{aligned} \tag{42}$$

Of course, the fixed weights P_{ij} can be removed from the summation inside the objective of the optimization problem above, without changing its optimal solution.

In Section 6, we present some numerical results obtained solving both Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$ through a modified version of the MATLAB function `fmmc.m` in the CVX package (http://cvxr.com/cvx/download/), which solves the SDP formulation of Problem FMMC presented in [5] and [9]. Such a modified version is basically obtained adding the linear term $\frac{\eta}{2}\sum_{i\neq j,\, i,j=1}^{n} P_{ij}$ inside the objective of the original optimization problem.

5 Theoretical Results for Problem FMMC-$l_0(\eta)$

In this section, we provide some theoretical results about the optimal solution of Problem FMMC-$l_0(\eta)$. They usually provide structural properties of the optimal solution similar to the ones obtained in Section 4 for Problem FMMC-$l_1(\eta)$, with some differences in the proofs. To differentiate the notation with respect to the one used for Problem FMMC-$l_1(\eta)$, we use the subscript "0" when referring to an optimal solution of Problem FMMC-$l_0(\eta)$.

a) Existence of an optimal solution.

The next result states the existence of an optimal solution to Problem FMMC-$l_0(\eta)$. Before stating it, we need to recall the following definition of lower semi-continuity, which is used in the proof of the next Proposition 5.

Definition 2. *Let $f : X \to \mathbb{R}$ be a real-valued function defined on a topological space X. Then, f is lower semi-continuous iff $\forall c \in \mathbb{R}$, the set $\{\underline{x} \in X : f(\underline{x}) \leq c\}$ is closed.*

Proposition 5. *Problem FMMC-$l_0(\eta)$ admits an optimal solution for every $\eta > 0$.*

Proof. The feasible set of Problem FMMC-$l_0(\eta)$ is convex, closed, and bounded, hence compact, as being $\mathbb{R}^m$ a finite-dimensional vector space. Moreover, its objective is lower semi-continuous since:

- the l_0-pseudo-norm regularization term $\|\underline{w}\|_0$ is lower semi-continuous (which is verified taking, for any $c > 0$, any convergent sequence $\{\underline{w}^{(k)}\}_{k=1}^{\infty}$ such that $\|\underline{w}^{(k)}\|_0 \leq c$ for all k, and observing that also its limit satisfies the same inequality);

- as it has been shown in the proof of Proposition 1, on the feasible set, the second-largest eigenvalue modulus $\mu(P(\underline{w}))$ is continuous (hence, also lower semi-continuous);
- the sum of two lower semi-continuous functions is lower semi-continuous, too.

Concluding, Problem FMMC-$l_0(\eta)$ involves the minimization of a lower semi-continuous objective function on a compact set, so an optimal solution to Problem FMMC-$l_0(\eta)$ exists by the generalized Weierstrass theorem [20, Theorem 2.43]. □

b) Effect of the regularization parameter.

Likewise Proposition 2, next Proposition 6 shows that the regularization parameter η has opposite effects on the two terms $\mu(\underline{w})$ and $\|\underline{w}\|_0$, when evaluated at an optimal solution. The proof is very similar to the one of Proposition 2, hence it is not reported.

Proposition 6. *Let $0 < \eta_1 < \eta_2$, and $\underline{w}_0^\circ(\eta_1), \underline{w}_0^\circ(\eta_2)$ be optimal solutions to Problem FMMC-$l_0(\eta_1)$ and Problem FMMC-$l_0(\eta_2)$, respectively. Then,*

i) $\mu(\underline{w}_0^\circ(\eta_1)) \leq \mu(\underline{w}_0^\circ(\eta_2))$;
ii) $\|\underline{w}_0^\circ(\eta_1)\|_0 \geq \|\underline{w}_0^\circ(\eta_2)\|_0$.

Differently from the case of Problem FMMC-$l_1(\eta)$ investigated in Section 4, however, also the sparsity

$$\begin{aligned} s(\underline{w}_0^\circ(\eta)) &:= 1 - \|\underline{w}_0^\circ(\eta)\|_0/m \\ &= \text{ fraction of zero components of } \underline{w}_0^\circ(\eta) \end{aligned}$$

is a monotonic function of η, as it is shown by the next proposition, whose proof is immediate and is, therefore, omitted.

Proposition 7. Let $0 < \eta_1 < \eta_2$, then $s(\underline{w}_0^\circ(\eta_1)) \leq s(\underline{w}_0^\circ(\eta_2))$.

Notice that, since

- *any subgraph with n vertices and $\hat{m} \leq m$ (non self-loop) non-zero weighted edges cannot be connected when $\hat{m} < n-1$ (as $n-1$ is the number of edges in a spanning tree);*
- *when the subgraph associated with a feasible $\underline{w}$ is disconnected, the corresponding $\mu(\underline{w})$ is equal to 1, so the optimal choice for $\underline{w}$ in Problem FMMC-$l_0(\eta)$ is $\underline{w} = \underline{0}_m$, when one limits to consider disconnected subgraphs;the value of the optimal sparsity for Problem FMMC-$l_0(\eta)$ cannot belong to the interval $\left(\frac{m-n+1}{m}, 1\right)$ (which corresponds with disconnected subgraphs with at least 1 non self-loop edge), whereas the value 1 is achievable, and it corresponds to the trivial case of a completely disconnected subgraph at optimality. Concluding, the possible values for the optimal sparsity for Problem FMMC-$l_0(\eta)$ are $0, \frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-n}{m}, \frac{m-n+1}{m}, 1$.*

The next result states the continuity of the optimal value of the objective in Problem FMMC-$l_0(\eta)$, and shows also that Problem FMMC-$l_0(\eta)$ "behaves" like Problem FMMC for η sufficiently small.

Proposition 8. *Let $\underline{w}_0^\circ(\eta)$ (resp., $\underline{w}^\circ$) be an optimal solution of Problem FMMC-$l_0(\eta)$ (resp., of Problem FMMC), and $\mu(\underline{w}_0^\circ(\eta))$ (resp., $\mu(\underline{w}^\circ)$) the value of the second-largest eigenvalue modulus of the associated weighted adjacency matrix $P(\underline{w}_0^\circ(\eta))$ (resp., $P(\underline{w}^\circ)$). Then,*

i) $\mu(\underline{w}_0^\circ(\eta)) + \eta\|\underline{w}_0^\circ(\eta)\|_0$ *depends continuously on η;*
ii) $\lim_{\eta\to 0^+} \mu(\underline{w}_0^\circ(\eta)) = \mu(\underline{w}^\circ)$;
iii) Given any sequence $\{\eta_k\}_{k=1}^{+\infty}$ convergent to 0, and an associated sequence $\{\underline{w}_0^\circ(\eta_k)\}_{k=1}^{+\infty}$, one can extract from the latter a subsequence that converges to an optimal solution of Problem FMMC.

Proof. i) Let $0 < \eta_1$, $0 < \eta_2$, and $\eta_1 \neq \eta_2$. Due to the optimality of $\underline{w}_0^\circ(\eta_1)$ and $\underline{w}_0^\circ(\eta_2)$ for Problems FMMC-$l_0(\eta_1)$ and FMMC-$l_0(\eta_2)$, resp., one gets

$$\mu(\underline{w}_0^\circ(\eta_1)) + \eta_1\|\underline{w}_0^\circ(\eta_1)\|_0 \leq \mu(\underline{w}_0^\circ(\eta_2)) + \eta_1\|\underline{w}_0^\circ(\eta_2)\|_0\,, \quad (43)$$

and

$$\mu(\underline{w}_0^\circ(\eta_2)) + \eta_2\|\underline{w}_0^\circ(\eta_2)\|_0 \leq \mu(\underline{w}_0^\circ(\eta_1)) + \eta_2\|\underline{w}_0^\circ(\eta_1)\|_0\,. \quad (44)$$

Then, combining (43) and (44), one obtains

$$\begin{aligned} &\mu(\underline{w}_0^\circ(\eta_1)) \\ &\leq \mu(\underline{w}_0^\circ(\eta_2)) + \eta_1\left(\|\underline{w}_0^\circ(\eta_2)\|_0 - \|\underline{w}_0^\circ(\eta_1)\|_0\right) \\ &\leq \mu(\underline{w}_0^\circ(\eta_1)) + \eta_2\left(\|\underline{w}_0^\circ(\eta_1)\|_0 - \|\underline{w}_0^\circ(\eta_2)\|_0\right) \\ &\quad + \eta_1\left(\|\underline{w}_0^\circ(\eta_2)\|_0 - \|\underline{w}_0^\circ(\eta_1)\|_0\right) \\ &= \mu(\underline{w}_0^\circ(\eta_1)) + (\eta_2 - \eta_1)\left(\|\underline{w}_0^\circ(\eta_1)\|_0 - \|\underline{w}_0^\circ(\eta_2)\|_0\right). \end{aligned} \quad (45)$$

Since $\underline{w}_0^\circ(\eta_1)$ and $\underline{w}_0^\circ(\eta_2)$ belong to the admissible set, which is compact, formula (45) implies (reversing also the roles of η_1 and η_2)

$$\begin{aligned} &\mu(\underline{w}_0^\circ(\eta_1)) + \eta_1\|\underline{w}_0^\circ(\eta_1)\|_0 \\ &\quad = \lim_{\eta_2\to\eta_1}\left(\mu(\underline{w}_0^\circ(\eta_2)) + \eta_1\|\underline{w}_0^\circ(\eta_2)\|_0\right) \\ &\quad = \lim_{\eta_2\to\eta_1}\left(\mu(\underline{w}_0^\circ(\eta_2)) + \eta_2\|\underline{w}_0^\circ(\eta_2)\|_0\right), \end{aligned} \quad (46)$$

from which item i) follows.

Item ii) is proved likewise item i), exploiting also the fact that $\underline{w}_0^\circ$ belongs to an admissible compact set, hence the term $\eta\|\underline{w}_0^\circ\|_0$ vanishes as η tends to 0 from the right.

Finally, item iii) is obtained combining item ii) with the compactness of the admissible set (which makes it possible to extract a convergent subsequence, starting from any subsequence belonging to that set) and the continuity of the second-largest eigenvalue modulus $\mu(P(\underline{w}))$ with respect to $\underline{w}$ (already shown in the proof of Proposition 1). □

Remark 7. A result similar to Proposition 8 holds also for Problem FMMC-$l_1(\eta)$. In that case, the proof of the corresponding item i) could be also obtained exploiting Berge's theorem of the maximum (see, e.g., [21, Section 3.3]).

c) Conditions under which $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_0(\eta)$.

The next result states conditions on the regularization parameter under which $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_0(\eta)$, or its unique optimal solution. An application of the result to the choice of the regularization parameter is given in Section 5 *f*).

Proposition 9. *Let $\eta \geq \frac{1}{n-1}$. Then $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_0(\eta)$. If $\eta > \frac{1}{n-1}$, then $\underline{w} = \underline{0}_m$ is its unique optimal solution.*

Proof. Let $\Delta\underline{w}$ be an arbitrary admissible increment of $\underline{w}$, starting from $\underline{w} = \underline{0}_m$. Then, the corresponding increment $\Delta f^{(0,\eta)}$ in the objective $f^{(0,\eta)}$ of Problem FMMC-$l_0(\eta)$ is

$$\Delta f^{(0,\eta)} = \mu(\Delta\underline{w}) - \mu(\underline{0}_m) + \eta\|\Delta\underline{w}\|_0 \,. \tag{47}$$

Now, one has $\mu(\underline{0}_m) = 1$ (as the associated weighted adjacency matrix is $P(\underline{0}_m) = I_n$). If $\|\Delta\underline{w}\|_0 < n-1$, then the graph associated with $\Delta\underline{w}$ is disconnected, and $\Delta f^{(0,\eta)} = \eta\|\Delta\underline{w}\|_0 \geq 0$. If $\|\Delta\underline{w}\|_0 \geq n-1$, then $\Delta f^{(0,\eta)} \geq \eta\|\Delta\underline{w}\|_0 - 1$, which is also non-negative by the assumption $\eta \geq \frac{1}{n-1}$. Then, if $\eta \geq \frac{1}{n-1}$, $\Delta f^{(0,\eta)}$ is non-negative for every arbitrary admissible $\Delta\underline{w} \neq \underline{0}_m$, hence $\underline{w} = \underline{0}_m$ is an optimal solution to Problem FMMC-$l_0(\eta)$. If $\eta > \frac{1}{n-1}$, then $\Delta f^{(0,\eta)}$ is positive for every arbitrary admissible $\Delta\underline{w} \neq \underline{0}_m$, hence $\underline{w} = \underline{0}_m$ is the only optimal solution to Problem FMMC-$l_0(\eta)$. □

The following example shows that the bound obtained in Proposition 9 is tight, at least if one does not impose further restrictions on the class of graphs to be considered.

Example 2. Likewise in Example 1, let $n = 2$ and $m = 1$. Then, again, the matrix $P(\underline{w})$ has the expression (29) and the eigenvalues 1 and $1 - 2w_1$. Hence, on the subset $(0, 1]$ of admissible solutions to Problem FMMC-$l_0(\eta)$, its objective is $|1 - 2w_1| + \eta$, whereas for $w_1 = 0$, the objective is equal to 1. Hence, $w_1 = 0$ is, respectively, the unique optimal solution to Problem FMMC-$l_0(\eta)$ for $\eta > 1$, one of its two optimal solutions for $\eta = 1$ (the other one being $w_1 = \frac{1}{2}$), and a suboptimal solution for $0 < \eta < 1$ (the optimal one being $w_1 = \frac{1}{2}$).

Remark 8. Interestingly, Example 2 demonstrates also that the functions $\mu(\underline{w}_0^\circ(\eta))$ and $\|\underline{w}_0^\circ(\eta)\|_0$ may be not continuous. Indeed, in this example, one has $\mu(\underline{w}_0^\circ(\eta)) = 0$ for $\eta \in (0, 1)$, and $\mu(\underline{w}_0^\circ(\eta)) = 1$ for $\eta > 1$. Moreover, $\|\underline{w}_0^\circ(\eta)\|_0 = 1$ for $\eta \in (0, 1)$, and $\|\underline{w}_0^\circ(\eta)\|_0 = 0$ for $\eta > 1$. A similar remark about the possible absence of continuity holds for Problem FMMC-$l_1(\eta)$, when one considers the same graph of this example.

d) Non differentiability of the objective at $\underline{w} = \underline{0}_m$.

Likewise Proposition 4, the next result shows that in general the objective of Problem FMMC-$l_0(\eta)$ is not differentiable at $\underline{w} = \underline{0}_m$.

Proposition 10. *Let $n \geq 2$. Then,*

- i) *the second-largest eigenvalue modulus term $\mu(\underline{w})$ is nondifferentiable at $\underline{w} = \underline{0}_m$ for $n \geq 3$, and differentiable for $n = 2$;*
- ii) *the l_0-pseudo-norm regularization term $\|\underline{w}\|_0$ is nondifferentiable at $\underline{w} = \underline{0}_m$;*
- iii) *the objective of Problem FMMC-$l_0(\eta)$ is nondifferentiable at $\underline{w} = \underline{0}_m$.*

Proof.

- i) Let us consider at first the case $n \geq 3$. Then, at $\underline{w} = \underline{0}_m$, proceeding likewise in the proof of Proposition 4, the directional derivative of the term $\mu(\underline{w})$ in the direction $\underline{e}_i$, i.e.,

$$\lim_{t \to 0^+} \frac{\mu(t\underline{e}_i) - \mu(\underline{0}_m)}{t} \,, \tag{48}$$

is equal to 0. Now, let $\underline{\hat{w}}$ be any admissible weight vector with $\underline{\hat{w}} > \underline{0}_m$ (elementwise). Then, at $\underline{w} = \underline{0}_m$, proceeding likewise in the proof of Proposition 4, the directional derivative of the term $\mu(\underline{w})$ in the direction $\underline{\hat{w}}$, i.e.,

$$\lim_{t \to 0^+} \frac{\mu(t\underline{\hat{w}}) - \mu(\underline{0}_m)}{t} \,, \tag{49}$$

is different from 0, and differs from

$$\sum_{i=1}^{m} \hat{w}_i \left(\lim_{t \to 0^+} \frac{\mu(t\underline{e}_i) - \mu(\underline{0}_m)}{t} \right) = 0 \tag{50}$$

(recall Remark 5). Then, we conclude that, for $n \geq 3$, the term $\mu(\underline{w})$ is nondifferentiable at $\underline{w} = \underline{0}_m$. Finally, for $n = 2$, the term $\mu(\underline{w})$ is differentiable, as shown by Example 2.

- ii) Nondifferentiability of the l_0-pseudo-norm term $\|\underline{w}\|_0$ at $\underline{w} = \underline{0}_m$ follows by the facts that $\|\underline{0}_m\|_0 = 0$, and in any neighborhood of $\underline{0}_m$ with radius $\varepsilon > 0$ there exists a feasible $\underline{w}^{(\varepsilon)}$ with $\|\underline{w}^{(\varepsilon)}\|_0 \geq 1$.
- iii) This follows combining the proofs of items i) and ii). □

e) Algorithmic issues.

The following result shows the combinatorial nature of Problem FMMC-$l_0(\eta)$, and also how solving an instance of such a problem can be reduced to solving several instances of (easier to solve) Problems FMMC. The upper bound $\lfloor \frac{1}{\eta} \rfloor$ of non self-loop edges in the next proposition comes from the consideration than any optimal solution $\underline{w}_0^\circ(\eta)$ of Problem FMMC-$l_0(\eta)$ cannot have more than $\lfloor \frac{1}{\eta} \rfloor$ non-zero components, otherwise the trivial choice $\underline{w} = \underline{0}_m$ is a better solution.

Proposition 11. *Any instance of Problem FMMC-$l_0(\eta)$ can be solved as follows:*

- *starting from the original graph $\mathcal{G}$ associated with the vertex-edge incidence matrix M, generate the set G_η of all its subgraphs $\mathcal{G}_k$ with n vertices and with at most $\lfloor \frac{1}{\eta} \rfloor$ non self-loop edges;*
- *for each subgraph $\mathcal{G}_k \in G_\eta$, find an optimal solution $\underline{w}^\circ(\mathcal{G}_k)$ of the instance of Problem FMMC associated with such subgraph, and the corresponding $\mu(\underline{w}^\circ(\mathcal{G}_k))$;*
- *find a subgraph $\mathcal{G}_{k^\circ} \in G_\eta$ that solves the optimization problem*

$$\underset{\mathcal{G}_k \in G_\eta}{\text{minimize}} \left(\mu(\underline{w}^\circ(\mathcal{G}_k)) + \eta\|\underline{w}^\circ(\mathcal{G}_k)\|_0 \right) . \tag{51}$$

Then,

- i) *the resulting vector $\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})$ is an optimal solution of Problem FMMC-$l_0(\eta)$;*
- ii) *all optimal solutions of Problem FMMC-$l_0(\eta)$ can be generated according to the procedure above. Moreover, the same optimal solution $\underline{w}_0^\circ(\eta)$ of Problem FMMC-$l_0(\eta)$ may be generated starting from more than one subgraph $\mathcal{G}_{k^\circ(\eta)} \in G_\eta$ (which is optimal for the*

optimization problem (51)), in case of non-uniqueness of the optimal $k^\circ(\eta)$. In particular, one of such subgraphs $\mathcal{G}_{k^\circ(\eta)}$ is the one $\mathcal{G}_{\tilde{k}(\eta)}$ whose non self-loop edges are the ones associated with the non-zero components of $\underline{w}_0^\circ(\eta)$.

Proof.

i) Since $\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})$ is feasible for Problem FMMC-$l_0(\eta)$, one gets, for any optimal solution $\underline{w}_0^\circ(\eta)$ of Problem FMMC-$l_0(\eta)$,

$$\begin{aligned}&\mu(\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})) + \eta\|\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})\|_0 \\ &\quad\geq \mu(\underline{w}_0^\circ(\eta)) + \eta\|\underline{w}_0^\circ(\eta)\|_0 \,.\end{aligned} \tag{52}$$

Let us show by contradiction that also the reverse inequality holds in (52), implying the equality therein. Suppose to the contrary that

$$\begin{aligned}&\mu(\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})) + \eta\|\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})\|_0 \\ &\quad> \mu(\underline{w}_0^\circ(\eta)) + \eta\|\underline{w}_0^\circ(\eta)\|_0 \,.\end{aligned} \tag{53}$$

Let $\mathcal{G}_{\tilde{k}(\eta)} \in G_\eta$ the subgraph whose non self-loop edges are the ones associated with the non-zero components of $\underline{w}_0^\circ(\eta)$. Hence, solving Problem FMMC on $\mathcal{G}_{\tilde{k}(\eta)}$, one obtains

$$\mu(\underline{w}^\circ(\mathcal{G}_{\tilde{k}(\eta)})) \leq \mu(\underline{w}_0^\circ(\eta))\,, \tag{54}$$

by the optimality of $\underline{w}^\circ(\mathcal{G}_{\tilde{k}(\eta)})$ for Problem FMMC on $\mathcal{G}_{\tilde{k}(\eta)}$, and the feasibility of $\underline{w}_0^\circ(\eta)$ for the same optimization problem, and

$$\|\underline{w}^\circ(\mathcal{G}_{\tilde{k}(\eta)})\|_0 \leq \|\underline{w}_0^\circ(\eta)\|_0\,, \tag{55}$$

by construction. Hence,

$$\begin{aligned}&\mu(\underline{w}^\circ(\mathcal{G}_{\tilde{k}(\eta)})) + \eta\|\underline{w}^\circ(\mathcal{G}_{\tilde{k}(\eta)})\|_0 \\ &\quad\leq \mu(\underline{w}_0^\circ(\eta)) + \eta\|\underline{w}_0^\circ(\eta)\|_0 \\ &\quad< \mu(\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})) + \eta\|\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})\|_0\,,\end{aligned} \tag{56}$$

which contradicts the optimality of $\mathcal{G}_{k^\circ(\eta)}$ for the optimization problem (51). Hence, (53) cannot hold, and $\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)})$ solves Problem FMMC-$l_0(\eta)$.

ii) Let us consider any optimal solution $\underline{w}_0^\circ(\eta)$ of Problem FMMC-$l_0(\eta)$. Likewise in the proof of item ii), considering the subgraph $\mathcal{G}_{\tilde{k}(\eta)} \in G_\eta$ whose non self-loop edges are the ones associated with the non-zero components of $\underline{w}_0^\circ(\eta)$, one obtains both (54) and (55) with the equality, due to the optimality of $\underline{w}_0^\circ(\eta)$ for Problem FMMC-$l_0(\eta)$. So, $\underline{w}_0^\circ(\eta)$ is also generated by the procedure detailed in the statement of the proposition. In particular, it is generated starting from the subgraph $\mathcal{G}_{\tilde{k}(\eta)}$. However, in general there may exist also other optimal subgraphs $\mathcal{G}_{k^\circ(\eta)}$ for the optimization problem (51), which generate the same $\underline{w}_0^\circ(\eta)$, i.e., such that $\underline{w}^\circ(\mathcal{G}_{k^\circ(\eta)}) = \underline{w}_0^\circ(\eta)$. □

Due to the combinatorial nature of Problem FMMC-$l_0(\eta)$, unfortunately, the number of instances of subproblems FMMC to be considered in formula (51) is in general very large (unless the original graph is "small", or η is large), however we remark that:

- each subproblem FMMC is convex and has a semidefinite programming formulation, which can be solved through the MATLAB function `fmmc.m` in the CVX package (http://cvxr.com/cvx/download/) already mentioned in Section 4;
- the number of subproblems FMMC to be considered depends on the parameter η, and is non-increasing with respect to η. In particular, larger values for η (which correspond with a larger desired sparsity) are associated with a smaller number of subproblems;
- some simplifications are possible, making it possible to reduce the number of subgraphs to be considered in formula (51). For instance, one can detect and remove all disconnected subgraphs (which can be detected, e.g., either checking the algebraic multiplicity of the associated Laplacian eigenvalue 1, or applying an algorithm presented in [22], which generates all connected subgraphs of a given graph, with the same number of vertices as in the original graph). Finally, isomorphic subgraphs could be also detected and represented by one single subgraph (although the approach should be in practice limited to subgraphs with a small number of edges, since the graph isomorphism problem belongs to the NP class of computational complexity [23]);
- Proposition 11 may suggest, as one possible heuristic to obtain good suboptimal solutions to an instance of Problem FMMC-$l_0(\eta)$, to generate a small number of random sparse subgraphs $\mathcal{G}_k$ of the original graph with vertex-edge incidence matrix M, solve the associated instances of Problem FMMC, then take, among the obtained optimal solutions $\underline{w}^\circ(\mathcal{G}_k)$, the one that minimizes $\mu(\underline{w}^\circ(\mathcal{G}_k)) + \eta\|\underline{w}^\circ(\mathcal{G}_k)\|_0$.

f) Choice of the regularization parameter.

Likewise in Section 4, the theoretical results presented above justify the following practical rule for choosing the regularization parameter η:

- given a positive integer N and a maximal acceptable increase $\varepsilon > 0$ for the second-largest eigenvalue modulus of P with respect to its optimal value μ_{FMMC}° in Problem FMMC, solve Problem FMMC-$l_0(\eta)$ in correspondence of N values $\eta^{(j)}$ for η such that
  - $0 < \eta^{(1)} < \eta^{(2)} < \cdots < \eta^{(N)} < \frac{1}{n-1}$, and
  - $\mu(\underline{w}_0^\circ(\eta^{(j)})) \leq \mu_{FMMC}^\circ + \varepsilon \; (j = 1, \ldots, N)$;
- choose $j^\circ \in \{1, \ldots, N\}$ that maximizes the sparsity $s(\underline{w}^\circ(\eta^{(j)}))$.

Remark 9. Differently from the case of Problem FMMC-$l_1(\eta)$ investigated in Section 4, a final reoptimization step is not needed after finding $\underline{w}^\circ(\eta^{(j^\circ)})$, since $\underline{w}^\circ(\eta^{(j^\circ)})$ solves an optimization problem including basically also the sparsity in its objective.

Finally, a possible way to choose the tolerance parameter ε (which has to be in any case smaller than $1 - \mu_{FMMC}^\circ$, again to avoid trivial optimal solutions) is given by formula (35), likewise in Section 4.

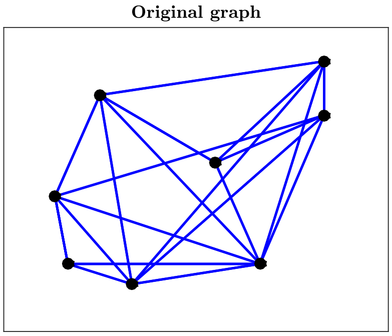

Fig. 1. A toy example modeled by a graph with eight vertices and 20 non self-loop edges.

g) Interpretation of Problem FMMC-$l_0(\eta)$ as a robust version of Problem FMMC.

Likewise Problem FMMC-$l_1(\eta)$, Problem FMMC-$l_0(\eta)$ has the following interpretation. Let us suppose that, for any given "nominal" choice of the weights w_i $(i = 1, \ldots, m)$, one has an "uncertainty" Δw_i such that $|\Delta w_i| \leq \delta U(w_i)$, for some fixed $\delta > 0$, where

$$U(w_i) := \begin{cases} 0 & \text{if } w_i = 0\,, \\ 1 & \text{if } w_i > 0\,. \end{cases} \tag{57}$$

Then, likewise in Section 4*f*), one can show that the second-largest eigenvalue modulus $\mu(\underline{w} + \Delta\underline{w})$ is bounded from above as

$$\mu(\underline{w} + \Delta\underline{w}) \leq \mu(\underline{w}) + 2\delta\|\underline{w}\|_0\,. \tag{58}$$

Then, an optimal "robust" choice of the nominal weight vector $\underline{w}$ is obtained minimizing the objective $\mu(\underline{w}) + 2\delta\|\underline{w}\|_0$ on the set of admissible weight vectors $\underline{w}$, and is obtained replacing the objective $\mu(\underline{w})$ of Problem FMMC with $\mu(\underline{w}) + 2\delta\|\underline{w}\|_0$. However, this is equivalent to solving Problem FMMC-$l_0((\eta)$ with the choice $\eta = 2\delta$. Likewise in Section 4*g*), when $\delta \geq 1$, $\underline{w} = \underline{0}_m$ is just an optimal nominal choice for the robust version of Problem FMMC just described, which is consistent with Proposition 9. In this case, however, that proposition shows that $\underline{w} = \underline{0}_m$ is an optimal nominal choice even under a less restrictive condition on δ: namely, for any $\delta \geq \frac{1}{2(n-1)}$.

6 Numerical Results

In this section, we first solve numerically Problems FMMC-$l_0(\eta)$ and FMMC-$l_1(\eta)$ on a toy example, in which both problems can be practically solved in a reasonably small amount of time, then their optimal solutions can be compared. Finally, at the end of the section, we compare the optimal solutions of Problems FMMC and FMMC-$l_1(\eta)$, when both problems are applied to a model of a wireless sensor network with a much larger number of nodes/edges.

6.1 Comparison of Problems FMMC-$l_1(\eta)$ and FMMC-$l_0(\eta)$

The comparison between the two sparse variations of Problem FMMC is performed on a graph with $n = 8$ vertices and $m = 20$ non self-loop edges, which is shown in Fig. 1. Problem FMMC-l_0 is solved by following the procedure described in Proposition 11. One can observe that, in this case, the number of all subgraphs with n vertices is equal to $2^{20} = 1{,}048{,}576$. However, since we are interested only in connected subgraphs, we first generate all such subgraphs (which have at least $n - 1 = 7$ non self-loop edges, since they must contain at least one spanning tree), then we associate all isomorphic connected subgraphs with a single representative connected subgraph. In this way, a total of 8,693 non isomorphic subgraphs is generated, on which Problem FMMC is solved, according to the procedure described in Proposition 11. The comparison between the optimal solutions to Problems FMMC-$l_1(\eta)$ and FMMC-$l_0(\eta)$ is performed by varying the regularization parameter η, and considering different ranges for such a parameter in the two problems, since equal values of the parameter are not directly comparable, as being associated with different regularizations. In particular, for both problems, we consider $N = 100$ different values for the regularization parameter equally spaced inside an interval I_1 for Problem FMMC-$l_1(\eta)$ and an interval I_0 for Problem FMMC-$l_0(\eta)$. From now on, we indicate with $\eta(l_1)$ the regularization parameter associated to Problem FMMC-$l_1(\eta)$, while $\eta(l_0)$ represents the regularization parameter associated to Problem FMMC-$l_0(\eta)$.

As a first step, we solve Problem FMMC on the graph shown in Fig. 1, obtaining its optimal solution $\underline{w}^\circ_{FMMC}$, whose second-largest eigenvalue modulus is equal to $\mu(\underline{w}^\circ_{FMMC}) = 0.3786$, and the sparsity is equal to $s(\underline{w}^\circ_{FMMC}) = 0$. Then, we study the optimal solutions achieved by Problem FMMC-$l_1(\eta)$ and Problem FMMC-$l_0(\eta)$. In practice, following the procedures illustrated in Sections 4*e*) and 5*f*), respectively, we aim at determining a feasible solution whose associated second-largest eigenvalue modulus is not much larger than its minimum possible value μ°_{FMMC}, and that, at the same time, provides a satisfactory sparsity. For both procedures, we choose $\rho = 1.5$, which is associated with the tolerance $\varepsilon = 0.145$, as $\mu^\circ_{FMMC} = 0.3786$ (see formula (35)). Hence, we are interested in studying how the optimal solutions to the two problems vary depending on $\eta(l_1)$ and $\eta(l_0)$, respectively, imposing the upper bound $\mu^\circ_{FMMC} + \varepsilon = 0.5236$ on μ. In particular, for Problem FMMC-$l_0(\eta)$ we consider 100 values of $\eta(l_0)$ equally spaced inside the interval $I_0 = [2 \cdot 10^{-8}, 0.08]$; while for Problem FMMC-$l_1(\eta)$ we consider 100 values for $\eta(l_1)$ equally spaced inside the interval $I_1 = [0.02, 0.198]$, since they provide comparable ranges of values for μ at optimality (for graphical reasons, the results in the next figures are reported at a lower resolution).

In Fig. 2, we report, as functions of the regularization parameter, the values of the second-largest eigenvalue modulus μ and the sparsity s for the optimal solutions of Problem FMMC-$l_1(\eta)$ (subplots (a) and (b), respectively and Problem FMMC-$l_0(\eta)$ (subplots (c) and (d), respectively). The results shown in Fig. 2 reveal, as expected, that Problem FMMC-$l_0(\eta)$ usually provides better solutions than Problem FMMC-$l_1(\eta)$. In fact, in the two respective ranges of values for the regularization parameter, the values of the second-largest eigenvalue modulus obtained solving Problem FMMC-$l_0(\eta)$ are comparable with the ones achieved solving Problem FMMC-$l_1(\eta)$, but with a better sparsity. However, from a computational point of view, solving Problem FMMC-$l_0(\eta)$ for the specific example takes a much longer

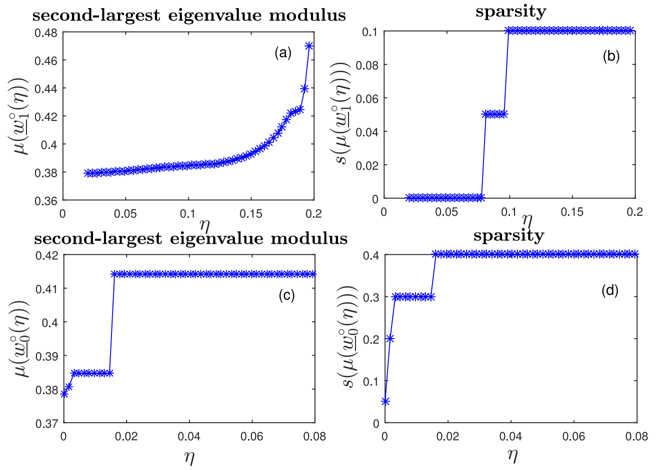

Fig. 2. Dependence on the regularization parameter of the second-largest eigenvalue modulus μ and the sparsity s for the optimal solutions of Problems FMMC-$l_1(\eta)$ (subplots (a) and (b)) and FMMC-$l_0(\eta)$ (subplots (c) and (d)).

time than solving Problem FMMC-$l_1(\eta)$ for the same example, i.e., about 40 seconds are needed to solve Problem FMMC-$l_1(\eta)$ for all the 100 values of its regularization parameter, whereas more than 4,000 seconds are required to solve Problem FMMC-$l_0(\eta)$ for all the 100 values of its regularization parameter, since this requires solving also all its subproblems FMMC (one time each). The numerical simulations have been performed using MATLAB R2015a on a notebook with a 1.60 GHz CPU and 8 GB of RAM.

In order to perform another comparison between the two approaches, we also proceed in the following way:

1) we fix a positive integer N_g, then we extract randomly N_g subgraphs over the total of $8{,}683$ non isomorphic connected subgraphs of the original graph. This number of subgraphs is chosen in order to be able to find an approximate solution to Problem FMMC-$l_0(\eta)$ in a time comparable to the one needed to solve Problem FMMC-$l_1(\eta)$ exactly (see the next step);
2) we apply a variation of the procedure described in Proposition 11, considering only the subgraphs generated in the step 1) above;
3) we repeat the two steps above for some number N_s of times;
4) we compute the average and standard deviation of the results obtained over the N_s repetitions.

In the following, for illustrative purposes, we always choose $N_g = 100$. We first consider the results achieved by the procedure described above when fixing $N_s = 1$. Fig. 3 shows the values of the second-largest eigenvalue modulus (subplot (a)) and of the sparsity for the suboptimal solution (subplot (b)) to Problem FMMC-$l_0(\eta)$ obtained in this case. Also in this case, the values of the sparsity obtained are better than the ones achieved solving Problem FMMC-$l_1(\eta)$, but larger values of the second-largest eigenvalue modulus are obtained compared with the exact solution of Problem FMMC-$l_0(\eta)$. In addition, when $\eta(l_0)$ is larger than 0.04, the obtained suboptimal solutions do not even satisfy the required constraint $\mu \leq 0.5236$.

We now consider the case $N_s = 10$. The plot on the top of Fig. 4 shows the average and standard deviation of the

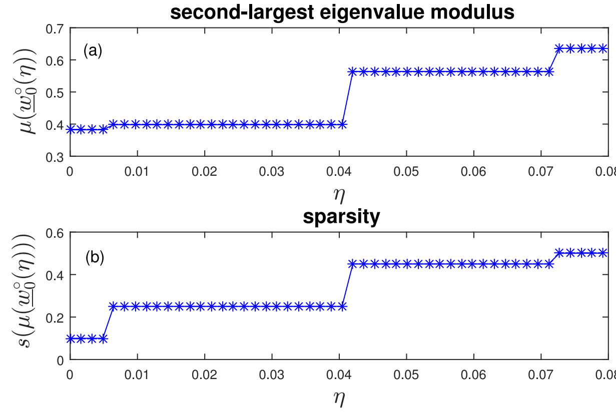

Fig. 3. Dependence on the regularization parameter of the second-largest eigenvalue modulus μ and the sparsity s for the suboptimal solution of Problem FMMC-$l_0(\eta)$, obtained when 100 subgraphs are randomly extracted from the whole set of connected non isomorphic subgraphs.

second-largest eigenvalue modulus of the suboptimal solution to Problem FMMC-$l_0(\eta)$ (subplot (a)), whereas subplot (b) does the same for the sparsity. Again, when $\eta(l_0)$ is larger than 0.06, in general the obtained suboptimal solutions do not even satisfy the required constraint $\mu \leq 0.5236$. In addition, due to the 10 repetitions, the time needed to obtain these results is about 10 times larger than the one needed to solve Problem FMMC-$l_1(\eta)$ exactly.

6.2 Comparison of Problems FMMC-$l_1(\eta)$, FMMC$_{\text{constr}}$-$l_1(\eta)$, and FMMC

In this part of the paper, we investigate numerically the optimal solutions of Problems FMMC-$l_1(\eta)$ and FMMC$_{\text{constr}}$-$l_1(\eta)$, comparing them with the one of Problem FMMC. In particular, as a test example, we consider a vertex-edge incidence matrix M corresponding to a model of a wireless sensor network with 50 vertices and 200 edges, generated in a similar way as the one in [9, Section 5.1]. The first two plots in Fig. 5 (subplots (a) and (b), respectively),

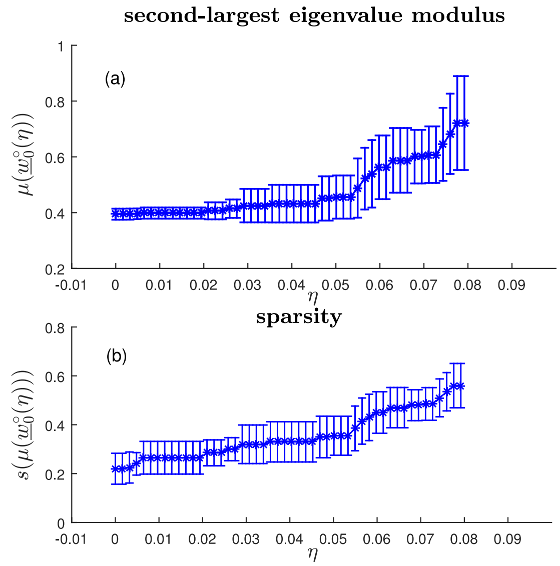

Fig. 4. Dependence on the regularization parameter of the average (over 10 trials) of the second-largest eigenvalue modulus μ (upper plot) and of the sparsity s (lower plot) for the suboptimal solution of Problem FMMC-$l_0(\eta)$, obtained when 100 subgraphs are randomly extracted from the whole set of connected non isomorphic subgraphs.

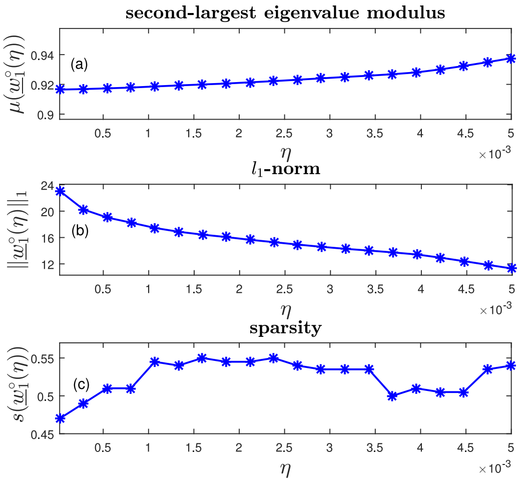

Fig. 5. Dependence on η of the second-largest eigenvalue modulus of the weighted adjacency matrix $P(\underline{w})$ evaluated at an optimal solution $\underline{w}_1^\circ(\eta)$ of Problem FMMC-$l_1(\eta)$, the l_1-norm of $\underline{w}_1^\circ(\eta)$, and its sparsity.

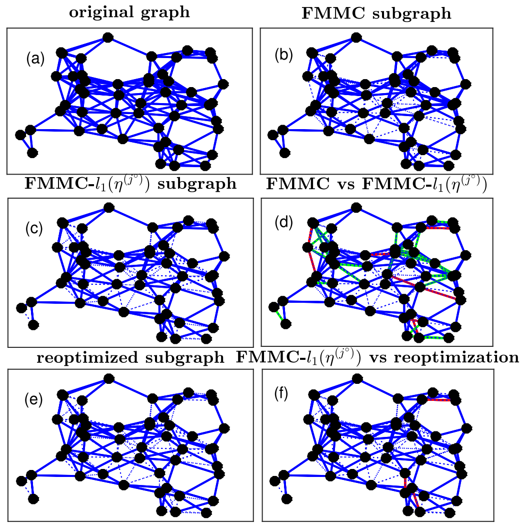

Fig. 6. A comparison of the subgraphs associated with non-zero weights in the optimal solutions to Problems FMMC and FMMC-$l_1(\eta^{(j^\circ)})$. See the main text for explanations about the colors used in the figure.

which refers to the behavior of an optimal solution $\underline{w}_1^\circ(\eta)$ with respect to η, confirm the statement of Proposition 2 about the opposite monotonic dependence on η of $\mu(\underline{w}_1^\circ(\eta))$ and $\|\underline{w}_1^\circ(\eta))\|_1$. Subplot (c) shows its sparsity $s(\underline{w}_1^\circ(\eta))$ as a function of η, which in this particular case is not a monotonic function of η. However, the plots also show that $\underline{w}_1^\circ(\eta)$ is sparser than the optimal solution of Problem FMMC, for all the considered values of η. So, they highlight the possibility, in this case, of finding a value of the parameter η for which the second-largest eigenvalue modulus $\mu(\underline{w}_1^\circ(\eta))$ is not much larger than its minimum possible value μ_{FMMC}°, and that, at the same time, provides a satisfactory sparsity of $\underline{w}_1^\circ(\eta)$. Again, in order to find such a parameter, we follow the procedure illustrated in Section 4e). We choose $\rho = 1.5$, associated with the tolerance $\varepsilon = 0.027$, as $\mu_{FMMC}^\circ = 0.9165$ in this particular case. We also consider $N = 20$ values $\eta^{(1)}, \ldots, \eta^{(N)}$ for the regularization parameter η (uniformly spaced in the interval $[2 \cdot 10^{-5}, 5 \cdot 10^{-3}]$, see Fig. 5), obtaining $j^\circ = 5$ and $\eta^{(j^\circ)} = 1.1 \cdot 10^{-3}$ as the optimal regularization parameter. For this value, we obtain $\mu(\underline{w}_1^\circ(\eta^{(j^\circ)})) = 0.9186$, $\|\underline{w}_1^\circ(\eta^{(j^\circ)})\|_1 = 17.45$, and $s(\underline{w}_1^\circ(\eta^{(j^\circ)})) = 0.545$. Compared with the optimal solution $\underline{w}_{FMMC}^\circ$ of Problem FMMC (for which $\mu(\underline{w}_{FMMC}^\circ) = 0.9165$, $\|\underline{w}_{FMMC}^\circ\|_1 = 23.71$, and $s(\underline{w}_{FMMC}^\circ) = 0.41$), the increase of the second-largest eigenvalue modulus, the decrease of the l_1-norm of the weight vector, and the increase of its sparsity are, respectively, about $0.2, 26$, and 25 percent. In terms of the mixing time (6), we obtain an increase of about 3 percent with respect to the value associated with $\underline{w}_{FMMC}^\circ$. Fig. 6 shows: the original graph associated with the given vertex-edge incidence matrix M (subplot (a)); its subgraph obtained keeping only the edges associated with non-zero weights in the optimal solution $\underline{w}_{FMMC}^\circ$ to Problem FMMC (subplot (b)); the one obtained keeping only the edges associated with the non-zero weights of $\underline{w}^\circ(\eta^{(j^\circ)})$ (subplot (c)); a comparison of the two subgraphs (subplot (d)), obtained merging such subgraphs and highlighting in blue the non-zero-weighted edges appearing in both graphs and in green (resp., red) the non-zero weighted edges of the optimal solution to Problem FMMC (resp., Problem FMMC-$l_1(\eta^{(j^\circ)})$) that are associated with zero weights in the optimal solution to Problem FMMC-$l_1(\eta^{(j^\circ)})$ (resp., Problem FMMC). In particular, starting from the original 200 edges joining different vertices, the optimal solution to Problem FMMC keeps 118 edges, while the optimal solution to Problem FMMC-$l_1(\eta^{(j^\circ)})$ keeps only 91 edges. The percentage of edge reduction when moving from $\underline{w}_{FMMC}^\circ$ to $\underline{w}_1^\circ(\eta^{(j^\circ)})$ is therefore about 23 percent.

As described in Section 4e), after finding the parameter $\eta^{(j^\circ)}$, an additional improvement may be obtained performing a reoptimization step, solving Problem FMMC on the sparser subgraph obtained deleting the edges associated with zero weights in the obtained optimal solution $\underline{w}_1^\circ(\eta^{(j^\circ)})$ to Problem FMMC-$l_1(\eta^{(j^\circ)})$. This step is illustrated in subplots (e) and (f) of Fig. 6, which shows in red the edges deleted by the reoptimization step. In this way, a new weight vector $\underline{w}_{\text{reopt}}^\circ$ is obtained with $\mu(\underline{w}_{\text{reopt}}^\circ) \leq \mu(\underline{w}_1^\circ(\eta^{(j^\circ)}))$ and $s(\underline{w}_{\text{reopt}}^\circ) \geq s(\underline{w}_1^\circ(\eta^{(j^\circ)}))$. So, compared with $\underline{w}_1^\circ(\eta^{(j^\circ)})$, the sparsity of the weight vector $\underline{w}_{\text{reopt}}^\circ$ either remains the same or even increases, whereas the second-largest eigenvalue modulus either remains the same or even decreases. Indeed, after the reoptimization step, we obtain $\mu(\underline{w}_{\text{reopt}}^\circ) = 0.9169$ and $s(\underline{w}_{\text{reopt}}^\circ) = 0.56$.

Finally, we report the results obtained solving Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ (in this case, for simplicity of comparison, for $\eta = \eta^{(j^\circ)}$), imposing the constraint that the 11 non self-loop edges associated with the vertex A in Fig. 7 are fixed, resp., to the values $0.1, 0.05, 0.25, 0.1, 0.01, 0.07, 0.05, 0.1, 0.02, 0.05, 0.1$, whose sum is $0.9 < 1$ (hence the problem is feasible). Since such constraints are not satisfied by $\underline{w}_1^{(j^\circ)}$, a significant change of the optimal solution is expected, with respect to the unconstrained version of the problem. Indeed, for the optimal solution $\underline{w}_{\text{free},1}^\circ(\eta^{(j^\circ)})$ to such an instance of Problem FMMC$_{\text{constr}}$-$l_1(\eta^{(j^\circ)})$, we obtain $\mu(\underline{w}_{\text{free},1}^\circ(\eta^{(j^\circ)})) = 0.9193$. Such an increase of $\mu(\underline{w}_{\text{free},1}^\circ(\eta^{(j^\circ)}))$ with respect to $\mu(\underline{w}_1^\circ(\eta^{(j^\circ)}))$ was also expected, as Problem FMMC$_{\text{constr}}$-$l_1(\eta^{(j^\circ)})$ is more constrained than Problem FMMC-$l_1(\eta^{(j^\circ)})$,

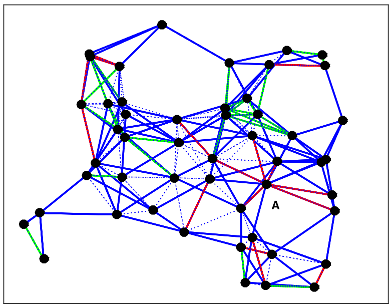

Fig. 7. A comparison of the subgraphs associated with non-zero weights in the optimal solutions to Problems FMMC-$l_1(\eta^{(j^\circ)})$ and FMMC$_{\text{constr}}$-$l_1(\eta^{(j^\circ)})$. Again, this is obtained by merging such subgraphs and highlighting in blue the non-zero-weighted edges appearing in both graphs and in green (resp., red) the non-zero weighted edges of the optimal solution to Problem FMMC$_{\text{constr}}$-$l_1(\eta^{(j^\circ)})$ that are associated with zero weights in the optimal solution to Problem FMMC-$l_1(\eta^{(j^\circ)})$ (resp., Problem FMMC$_{\text{constr}}$-$l_1(\eta^{(j^\circ)})$).

and has the same objective (including in the objective also the fixed weights).

7 Conclusions

In the paper, we have presented some theoretical and numerical results about several sparse variations of the Fastest Mixing Markov-Chain problem. Among possible future developments, we mention:

- the possibility of using other sparsity-enforcing regularization terms (such as the reweighted l_1-norm [24], the group LASSO [25] and the sparse group LASSO [26]);
- the possible extension of the nondifferentiability results provided by Propositions 4 and 10 to other values for the weight vector $\underline{w}$. This could be useful to motivate the choice of suitable alternative iterative algorithms to solve Problems FMMC-$l_1(\eta)$ and FMMC-$l_0(\eta)$;
- an investigation of theoretical bounds on the degree of sub-optimality in sparsity of the optimal solution to Problem FMMC$_{\text{constr}}$-$l_1(\eta)$ with respect to the one achieved using the l_0-pseudo-norm;
- the possibility to obtain probabilistic guarantees on the heuristic proposed in Section 5e) to solve Problem FMMC$_{\text{constr}}$-$l_0(\eta)$ suboptimally, using the scenario optimization approach [27];
- the possibility of solving the proposed regularized optimization problems in a distributed way;
- an extension of the theoretical analysis to nonlinear and stochastic consensus problems.

References

[1] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2010.

[2] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, “Control of complex networks,” *Nature*, vol. 473, pp. 167–173, 2011.

[3] E. Lovisari and S. Zampieri, “Performance metrics in the average consensus problem: A tutorial,” *Annu. Rev. Control*, vol. 36, pp. 26–41, 2012.

[4] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” in *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[5] S. Boyd, P. Diaconis, and L. Xiao, “Fastest mixing Markov chain on a graph,” *SIAM Rev.*, vol. 46, pp. 667–689, 2004.

[6] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, “Communication constraints in the average consensus problem,” *Automatica*, vol. 44, pp. 671–684, 2008.

[7] J. Batson, A. Spielman, N. Srivastava, and S.-H. Teng, “Spectral sparsification of graphs: Theory and algorithms,” *Commun. ACM*, vol. 56, pp. 87–94, 2013.

[8] R. Tibshirani, “Regression shrinkage and selection via the LASSO,” *J. Roy. Statistical Soc., Series B*, vol. 58, pp. 267–288, 1994.

[9] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” *Syst. Control Lett.*, vol. 53, pp. 65–78, 2004.

[10] N. Dhingra, F. Lin, M. Fardad, and M. R. Jovanović, “On identifying sparse representations of consensus networks,” in *Proc. 3rd IFAC Workshop Distrib. Estimation Control Netw. Syst.*, 2012, vol. 3, part 1, pp. 305–310.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, pp. 1–122, 2011.

[12] G. Gnecco, R. Morisi, and A. Bemporad, “Sparse solutions to the average consensus problem via l_1-norm regularization of the fastest mixing Markov-chain problem,” in *Proc. 53rd IEEE Int. Conf. Decision Control*, 2014, pp. 2228–2233.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] D. Spielman, “Spectral graph theory,” in *Combinatorial Scientific Computing*, U. Naumann and O. Schenk, Eds. Boca Raton, FL, USA: CRC Press, 2012, ch. 18, pp. 495–524.

[16] H. Dym, *Linear Algebra in Action*. Providence, RI, USA: AMS, 2007.

[17] T. Hastie, J. Taylor, and R. Tibshirani, “Forward stagewise regression and the monotone LASSO,” *Electron. J. Statist.*, vol. 1, pp. 1–29, 2007.

[18] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[19] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[20] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*. New York, NY, USA: Springer-Verlag, 2006.

[21] N. L. Stokey, R. E. Lucas, and E. C. Prescott, *Recursive Methods in Economic Dynamics*. Cambridge, MA, USA: Harvard Univ. Press, 1989.

[22] G. Rúcker, and C. Rúcker, “Automatic enumeation of all connected subgraphs,” *MATCH Commun. Math. Comput. Chemistry*, vol. 41, pp. 145–149, 2000.

[23] M. R. Garey and D. S. Johson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.

[24] E. J. Candès, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted l_1 minimization,” *J. Fourier Anal. Appl.*, vol. 14, pp. 877–905, 2008.

[25] E. Yuan and Y. Lin, “Model selection and estimation in regression with grouped variables,” *J. Roy. Statistical Soc., Series B*, vol. 68, pp. 49–67, 2007.

[26] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, “A sparse group LASSO,” *J. Comput. Graphical Statist.*, vol. 22, pp. 231–245, 2013.

[27] M. C. Campi and S. Garatti, “The exact feasibility of randomized solutions of uncertain convex programs,” *SIAM J. Optimization*, vol. 19, pp. 1211–1230, 2008.

Giorgio Gnecco received the “Laurea” (MSc) degree cum laude in telecommunications engineering in 2004 and the PhD degree in mathematics and applications in 2009, both from the University of Genoa, Italy. After being a postdoctoral researcher at the University of Genoa from 2009 to 2013, from 2013, he is an assistant professor in control systems at the IMT—Institute for Advanced Studies, Lucca, Italy. He has been an associate editor for the *IEEE Transactions on Neural Networks and Learning Systems* since 2013. His current research topics include optimization, machine learning, and their application to control systems.

Rita Morisi received the “Laurea” (MSc) degree cum laude in mathematics from the University of Bologna, Italy, in 2011. She is currently working toward the PhD degree at IMT—Institute for Advanced Studies, Lucca, Italy, in computer decision and systems science. Her research is mainly focused on the development and study of machine learning methods. In particular, her work examines, from both a theoretical and practical point of view, the application of graph-based techniques in machine learning and in control theory.

Alberto Bemporad received the master’s degree in electrical engineering in 1993 and the PhD degree in control engineering in 1997 from the University of Florence, Italy. He spent the 1996/1997 academic year at the Center for Robotics and Automation, Department of Systems Science & Mathematics, Washington University, St. Louis, as a visiting researcher. In 1997-1999, he held a postdoctoral position at the Automatic Control Laboratory, ETH Zurich, Switzerland, where he collaborated as a senior researcher in 2000-2002. In 1999-2009, he was with the Department of Information Engineering, University of Siena, Italy, becoming an associate professor in 2005. In 2010-2011, he was with the Department of Mechanical and Structural Engineering, University of Trento, Italy. In 2011, he joined the IMT Institute for Advanced Studies Lucca, Italy, as a full professor, where he also became the director in 2012. In 2011, he cofounded the company ODYS S.r.l. He has published more than 250 papers in the areas of model predictive control, hybrid systems, automotive control, multiparametric optimization, computational geometry, robotics, and finance. He is author or coauthor of various MATLAB toolboxes for model predictive control design, including the Model Predictive Control Toolbox (The Mathworks, Inc.). He was an associate editor of the *IEEE Transactions on Automatic Control* during 2001-2004 and the chair in the Technical Committee on Hybrid Systems of the IEEE Control Systems Society in 2002-2010. He received the IFAC High-Impact Paper Award for the 2011-2014 triennial. He has been an IEEE Fellow since 2010.