

A Joint Markov Model for Communities, Connectivity and Signals defined over Graphs

Stefania Colonnese *Senior Member, IEEE*, Paolo Di Lorenzo *Senior Member, IEEE*,
Tiziana Cattai, Gaetano Scarano, Fabrizio De Vico Fallani

Abstract—Real-world networks are typically described in terms of nodes, links, and communities, having signal values often associated with them. The aim of this paper is to introduce a novel Compound Markov random field model (Compound MRF, or CMRF) for signals defined over graphs, encompassing jointly signal values at nodes, edge weights, and community labels. The proposed CMRF generalizes Markovian models previously proposed in the literature, since it accounts for different kinds of interactions between communities and signal smoothness constraints. Finally, the proposed approach is applied to (joint) graph learning and signal recovery. Numerical results on synthetic and real data illustrate the competitive performance of our method with respect to other state-of-the-art approaches.

Index Terms—Graph Signal Processing, Markov random field, graph community, graph learning, graph signal denoising.

I. INTRODUCTION

SIGNALS on graphs (SoG) describe a huge variety of complex systems like social, biological, and technological networks [1]. Several SoG models have been proposed in the literature to, e.g., infer network topology, denoise measurements, predict missing data, etc. [2]–[4]. Among others, Markovian models are widely adopted to model complex systems involving several random variables; such models are useful whenever global constraints on single random variables can be expressed exactly, or approximately, in terms of local constraints. Fruitful applications of Markovian models can be found in signal and image processing [5]–[9], and, recently, they have been applied also to graphs [11], [12], especially for graph inference purposes. In particular, SoG are often modeled as Gaussian Markov Random Fields (GMRF), having a precision matrix related to the Laplacian matrix of the underlying graph, see, e.g., [13], [15]. Also, in [14], the authors infer the graph topology based on the potential energy of the associated MRF. In [16], SoG are modeled as an irregular sampling of a continuous space GMRF. Finally, in [17] the authors proposed an MRF to model edges and signals observed over graphs in a joint fashion.

Contribution. Many social or biological networks are organized into communities, i.e., subsets of nodes that feature peculiar interconnections among them. Clearly, incorporating such side information in SoG modeling may improve the performance of several learning tasks. In this paper, we address

the joint modeling of the signal values, edges connectivity, and the underlying community structure using distinct but related random fields. We introduce a novel compound MRF (CMRF) model built by the signal, edge weight and community labeling fields. To the best of our knowledge, such holistic modeling has never been proposed up to date. The proposed CMRF generalizes Markovian models in the literature and it can be tuned to represent different kind of interactions between graph communities, such as attractive or repulsive behavior between nodes. Also, our CMRF model paves the way for designing joint solutions of GSP problems such as graph learning, community detection, and signal recovery, which are conventionally independently tackled. The proposed CMRF is applied to the recovery of the graph topology and signal from data; numerical simulations illustrate its advantages with respect to state-of-the-art competitors.

Outline. The structure of the paper is as follows. In Sec. II we introduce the proposed CMRF. In Sec. III we apply the CMRF to graph learning and signal recovery, and in Sec. IV we evaluate the performance of our approach against other techniques. Finally, Sec. V draws some conclusions.

II. A COMPOUND MARKOV RANDOM FIELD MODEL FOR SIGNALS ON GRAPHS

In this section, we introduce the proposed CMRF model, which accounts jointly for signal values observed at nodes, graph connectivity, and the underlying community structure. Let us consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} denotes the set of N vertices, and $\mathcal{E} = \{e_{ij} = (v_i, v_j)\}$, for all $v_i, v_j \in \mathcal{V}$, represents the set of edges. A real weight w_{ij} is associated to each edge $e_{ij} \in \mathcal{E}$. Let $\mathbf{W} = \{w_{ij}\}_{i,j=1}^N$ be the $N \times N$ real matrix collecting all edge weights. Also, let \mathbf{A} be the $N \times N$ binary adjacency matrix, whose elements are such that $a_{ij} = 1$ if $w_{ij} \neq 0$, and 0 otherwise. The graph Laplacian is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where \mathbf{D} is the diagonal degree matrix, having $d_{ii} = \sum_{j=1}^N a_{ij}$, for all $i = 1, \dots, N$.

An M -dimensional signal $\mathbf{x}(v) \in \mathbb{R}^M$ is associated to vertex v , for all $v \in \mathcal{V}$. Furthermore, we consider a partition of \mathcal{V} into C disjoint communities, assigning each vertex to a community using the binary vector $\boldsymbol{\gamma}(v) = [\gamma_1(v), \dots, \gamma_C(v)]^T$, for all $v \in \mathcal{V}$, where $\gamma_c(v) = 1$ if node v belongs to community c , and zero otherwise. Thus, inspired by the *pixel-and-edge* model [18], [19], adopted in image processing to jointly represent smooth regions and their boundaries [20], we represent the multidimensional signal $\mathbf{x}(v)$, $v \in \mathcal{V}$, the graph adjacency $a(e)$, $e \in \mathcal{E}$, and the community structure $\boldsymbol{\gamma}(v)$, $v \in \mathcal{V}$, as realizations of the following triplet of random fields: i) *Signal*

S. Colonnese, P. Di Lorenzo, T. Cattai, and G. Scarano are with the DIET Dept., University of Rome "Sapienza", Rome, Italy. T. Cattai and Fabrizio De Vico Fallani are with Inria, Aramis project-team, F-75013, Paris, France and Institut du Cerveau et de la Moelle pinire, ICM, F-75013, Inserm, U 1127, F-75013, CNRS, UMR 7225, F-75013, Sorbonne Universit, F-75013, Paris, France. FD acknowledges support from the Agence Nationale de la Recherche through contract number ANR15-NEUC-0006-02.

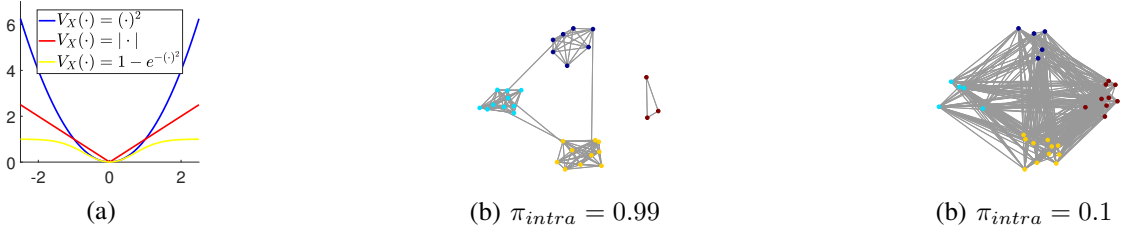


Fig. 1. (a) Clique potential functions $V_{\mathcal{X}}(\cdot)$ (the case $V_{\mathcal{X}}(\cdot) = (\cdot)^2$ corresponds to GMRF); (b) assortative and (c) disassortative community structure.

field $\mathcal{X} : \mathcal{V} \rightarrow \mathbb{R}^M$; ii) Edge field $\mathcal{A} : \mathcal{E} \rightarrow \{0, 1\}$; iii) Community field $\Gamma : \mathcal{V} \rightarrow \mathcal{B} = \{\mathbf{1}_i, i = 1, \dots, C\}$, where \mathcal{B} represents the set of unit-norm vectors $\mathbf{1}_i \in \{0, 1\}^C$ of the M -dimensional standard basis. Then, we define the *compound domain* $\mathcal{U} \triangleq \mathcal{V} \cup \mathcal{E} \cup \mathcal{V}$, and the related *compound random field*: $\tilde{X} : \mathcal{U} \rightarrow \mathbb{R}^M \cup \{0, 1\} \cup \mathcal{B}$, being:

$$\tilde{x}(u) = \begin{cases} \mathbf{x}_v \in \mathbb{R}^M \\ a_e \in \{0, 1\} \\ \gamma_v \in \mathcal{B}, \end{cases} \quad \text{for } u = [v, e, v] \in \mathcal{U}.$$

If \tilde{X} is an MRF, any u is associated to a neighborhood system \mathcal{N}_u , and the local Markov property holds, i.e. $p(\tilde{x}_u | \mathcal{U}) = p(\tilde{x}_u | \mathcal{N}_u)$. Also, a MRF obeys a Gibbs distribution [21]:

$$p_{\tilde{X}}(\tilde{\mathbf{x}}) = \frac{\exp\{-U(\tilde{\mathbf{x}})\}}{\sum_{\tilde{\mathbf{x}} \in \tilde{\mathcal{X}}} \exp\{-U(\tilde{\mathbf{x}})\}} \quad (1)$$

where $U(\tilde{\mathbf{x}}) = \sum_u \sum_{u' \in \mathcal{N}(u)} V(\|\tilde{x}_u - \tilde{x}_{u'}\|)$ is a global potential function, with $V(\cdot)$ denoting a non-negative non-decreasing function computed on cliques¹, i.e., the clique potential. In particular, for the proposed CMRF \tilde{X} , the Gibbs distribution (1) depends on the component fields $\mathcal{X}, \mathcal{A}, \Gamma$, as:

$$p_{\mathcal{X}, \mathcal{A}, \Gamma}(\mathbf{x}, a, \gamma) = \frac{\exp\{-U(\mathbf{x}, a, \gamma)\}}{Z}, \quad (2)$$

where the constant term Z is set as $Z = \sum_{\tilde{\mathbf{x}} \in \tilde{\mathcal{X}}} \exp\{-U(\tilde{\mathbf{x}})\}$. In the *pixel-and-edge* MRF used in image processing applications [19], the MRF jointly models the image luminance values and the discontinuities along smooth regions boundaries. This is achieved by factorizing the Gibbs distribution into two terms: one representing the relationship between image values and edges, and the other modeling intrinsic image edge properties, e.g., the spatial continuity. Herein, we follow an analogous approach, thus separately modeling three kinds of relationships, namely: i) signal-edge interactions, ii) community-edge interactions, and iii) intrinsic edge constraints. To this aim, the Gibbs distribution in (2) is factorized into the product of three terms, and the potential function $U(\mathbf{x}, a, \gamma)$ is cast as:

$$U(\mathbf{x}, a, \gamma) = \underbrace{\sum_{i \in \mathcal{V}, j \in \mathcal{N}_i} a_{ij} V_{\mathcal{X}}(\|\mathbf{x}_i - \mathbf{x}_j\|)}_{U_{\mathcal{X}}(\mathbf{x}, a)} + \underbrace{\sum_{i \in \mathcal{V}, j \in \mathcal{N}_i} a_{ij} V_{\Gamma}(\|\gamma_i - \gamma_j\|)}_{U_{\Gamma}(a, \gamma)} + \underbrace{\sum_{e \in \mathcal{E}, k \in \mathcal{N}_e} V_{\mathcal{A}}(a, a_k)}_{U_{\mathcal{A}}(a)} \quad (3)$$

¹A clique is a set of fully connected elements.

being $V_{\mathcal{X}}, V_{\Gamma}, V_{\mathcal{A}}$ three non-negative potential functions. The three terms in the potential function $U(\mathbf{x}, a, \gamma)$ have the following meaning. The term $U_{\mathcal{X}}(\mathbf{x}, a)$ is jointly related to the signal and the edge processes: it enforces signal smoothness on pairwise connected nodes and, vice-versa, connectivity over nodes with similar signal values. The term $U_{\Gamma}(a, \gamma)$ controls the probability of observing edges between different communities. Finally, the term $U_{\mathcal{A}}(a)$ implies constraints on edges stemming from the same node.

Choice of the potential functions. For simplicity, but without loss of generality, let us consider the case of scalar SoG (i.e., $M = 1$), which is commonly used in the literature. Regarding $V_{\mathcal{X}}$, the choice of a quadratic prior as $V_{\mathcal{X}}(\cdot) = (\cdot)^2$ corresponds to using a GMRF model [13]. However, non-quadratic priors such as $V_{\mathcal{X}}(\cdot) = |\cdot|$ or $V_{\mathcal{X}}(\cdot) = 1 - e^{-(\cdot)^2}$ can be used to limit the penalty computed on largely different signal values (see, e.g., Fig.1(a)), and are well suited to model SoGs where a few nodes largely differ from their neighborhood as, e.g., in the case of sparse or spiky signals.

Regarding V_{Γ} , the distance metric should assert if two nodes belong or not to the same community. Then, the potential function V_{Γ} may take the form $V_{\Gamma}(\cdot) = 2 - \|\cdot\|_0$ or $V_{\Gamma}(\cdot) = \|\cdot\|_0$, depending on whether high or low potential energy wants to be assigned to links within the same community. Thereby, V_{Γ} can model different network structures such as, e.g., disassortative or assortative ones (see, e.g., Fig.1 (b) and (c)). The further prior $U_{\mathcal{A}}(a)$ in (3) accounts for relative edge positioning and can be used to model SoGs with constrained graphs layouts. A relevant example is that of time-variant SoGs, where the vertex-time domain may obey to specific topological constraints, see, e.g., [22].

On the structure of $U_{\mathcal{X}}$. As previously mentioned, a GMRF is a particular case of our framework, corresponding to the selection of a quadratic cost function $U_{\mathcal{X}}$, i.e., $U_{\mathcal{X}}(\mathbf{x}, a) = \frac{1}{2} \mathbf{x}^T \mathbf{L} \mathbf{x}$. For general choices of $U_{\mathcal{X}}$, a more general expression can be found for the proposed CMRF. In particular, let us consider the Taylor series expansion of $V_{\mathcal{X}}$ up to order \bar{K} , i.e., $V_{\mathcal{X}}(x_i - x_j) \approx \sum_{k=1}^{\bar{K}} \frac{1}{k!} V_{\mathcal{X}}^{(k)}(0) (x_i - x_j)^k$, where $V_{\mathcal{X}}^{(k)}(0)$ is the k -th order derivative of $V_{\mathcal{X}}$ evaluated at 0. Then, $U_{\mathcal{X}}(\mathbf{x}, a)$ in (3) writes as: $U_{\mathcal{X}}(\mathbf{x}, a) = \sum_{k=1}^{\bar{K}} \frac{1}{k!} V_{\mathcal{X}}^{(k)}(0) \sum_{i \in \mathcal{V}, j \in \mathcal{N}(i)} a_{ij} (x_i - x_j)^k$. Introducing now the vector $\mathbf{h}^{(k)}(x)$ given by

$$\mathbf{h}^{(k)}(\mathbf{x}) = [\underbrace{1 \cdots 1}_{0\text{-th}} \underbrace{x_0 \cdots x_{N-1}}_{1\text{-st}} \cdots \underbrace{x_0^k \cdots x_{N-1}^k}_{k\text{-th order}}],$$

which collects the Hadamard (entry-wise) powers of the

$[x_0 \cdots x_{N-1}]$ up to the k -th order, and leveraging $(a-b)^k = \sum_{l=0}^k \binom{k}{l} (-1)^l a^l b^{k-l}$, we obtain:

$$U_{\mathcal{X}}(\mathbf{x}, a) = \sum_{k=1}^{\bar{K}} \mathbf{h}^{(k)}(\mathbf{x})^T \mathcal{A}^{(k)} \mathbf{h}^{(k)}(\mathbf{x}), \quad (4)$$

where $\mathcal{A}^{(k)} = (\mathbf{V}^{(k)} \otimes \mathbf{A}) \mathbf{J}$, with $\mathbf{V}^{(k)}$ denoting a diagonal matrix whose l -th entry is equal to $\frac{V_{\mathcal{X}}^{(k)}(0)}{k!} \binom{k}{l} (-1)^l$; also, \otimes denotes the Kronecker product, and \mathbf{J} is the exchange matrix [25]. Eq.(4), which for symmetric potential function and $\bar{K} = 2$, boils down to the GMRF quadratic form $U_{\mathcal{X}}(\mathbf{x}, a) = \mathbf{h}^{(2)}(\mathbf{x})^T (\mathbf{V}^{(2)} \otimes \mathbf{A}) \mathbf{J} \mathbf{h}^{(2)}(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{L} \mathbf{x}$, rephrases the potential Eq.(4), as an element of the ring of polynomials. This model can be leveraged for numerical and analytical solutions of graph inference problems. This is left for further studies.

III. APPLICATION TO GSP INFERENCE PROBLEMS

In this section, we leverage the CMRF in two fundamental GSP tasks: i) graph learning; ii) and joint signal and graph topology recovery, when the communities are known but the actual links are not. This occurs for instance, in social networks, where users are profiled as belonging to a community, e.g. by age, school or association, but mutual interactions and correlation between users have to be estimated to setup personalized advertising [23]. In such a case, the community information can be leveraged in estimating connectivity pattern between users' choices. Likewise, in brain networks, the communities represent areas which are known to be involved in given tasks [24], but the underlying graph topology has to be estimated. These problems can be tackled leveraging the proposed CMRF, as we will illustrate next.

1) *Graph Learning*: Let us assume that we know the signal \mathbf{x}_i , $i = 1, \dots, N$, and the community partition γ_i , $i = 1, \dots, C$. Then, we perform Maximum a Posteriori (MAP) estimation seeking for the edges $\{a_{ij}\}$ that maximize the Gibbs distribution in (2). Exploiting (3), this task coincides with the following minimization problem:

$$\begin{aligned} \min_{\mathbf{A}} \sum_{i \in \mathcal{V}, j \in \mathcal{N}(i)} a_{ij} \left(V_{\mathcal{X}}(\|\mathbf{x}_i - \mathbf{x}_j\|) + V_{\Gamma}(\|\gamma_i - \gamma_j\|) \right) \\ \text{subject to } \|\mathbf{A}\|_0 = 2K, \end{aligned} \quad (5)$$

where K represents the number of edges assumed to be in the graph. Since the objective function is linear in $\{a_{ij}\}$, if $V_{\mathcal{X}}(\cdot)$ is a non-negative and non-decreasing function, the solution of (5) coincides with the selection of the K edges corresponding to the K smallest values of $V_{\mathcal{X}}(\|\mathbf{x}_i - \mathbf{x}_j\|) + V_{\Gamma}(\|\gamma_i - \gamma_j\|)$, for all $i, j = 1, \dots, N$. When no prior on the community field is available, the term $V_{\Gamma}(\cdot)$ vanishes, and the MAP solution is driven only by the component $V_{\mathcal{X}}(\|\mathbf{x}_i - \mathbf{x}_j\|)$. Interestingly, in such a case, relaxing the discrete coefficients $\{a_{ij}\}$ to be real variables in the interval $[0, 1]$, and considering $V_{\mathcal{X}}(\|\mathbf{x}_i - \mathbf{x}_j\|) = 1 - \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, problem (5) becomes: $\max_{a_{ij} \in [0, 1]} \sum_{ij} a_{ij} \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, which leads to the so-called Gaussian weights [16], i.e., $a_{ij} = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2)$. Thus, the CMRF extends to real weights, and the graph learning in (5) includes as particular cases

Algorithm 1 : Joint Signal and Graph Topology Recovery

Data: \mathbf{y} , γ . Set $\mathbf{x}(0) = \mathbf{y}$; for $t \geq 0$, repeat:

Step 1: Solve the SCA optimization problem

$$\begin{aligned} (\hat{\mathbf{x}}(t), \mathbf{A}(t+1)) = \arg \min_{\mathbf{x}, \mathbf{A}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2 \\ + \lambda_1 \cdot \tilde{U}_{\mathcal{X}, \mathbf{A}}(\mathbf{x}, \mathbf{A}; \mathbf{x}(t), \mathbf{A}(t)) + \lambda_2 \cdot U_{\Gamma}(\mathbf{A}, \gamma) \\ \text{subject to } \|\mathbf{A}\|_0 = 2K, \end{aligned} \quad (7)$$

Step 2: Compute: $\mathbf{x}(t+1) = \mathbf{x}(t) + \alpha(t)(\hat{\mathbf{x}}(t) - \mathbf{x}(t))$

Gaussian distance weights [16], covariance based weights [22], Intensity/Distance based weights [14], among the others.

Finally, we highlight that our learning strategy in (5) might also be modified to select edges whose potential energy falls below a given threshold θ . Either K_0 in (5) or θ shall be properly tuned given the system constraints (e.g., number of neighbors per node, maximum number of links, etc.).

2) *Joint Topology and Signal Recovery*: In this task, we aim to jointly recover the graph signal and its topology from noisy observations given by: $\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i$, $i = 1, \dots, N$, where $\mathbf{v} = [\mathbf{v}_1^T, \dots, \mathbf{v}_N^T]$ is a zero-mean noise vector with covariance matrix $\sigma^2 \mathbf{I}$. Let $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_N^T]$ and $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]$. Then, exploiting the proposed CMRF model in (3), the task of joint recovery of \mathbf{x} and \mathbf{A} can be cast as:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{A}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2 + \lambda_1 \cdot U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}, \mathbf{A}) + \lambda_2 \cdot U_{\Gamma}(\mathbf{A}, \gamma) \\ \text{subject to } \|\mathbf{A}\|_0 = 2K, \end{aligned} \quad (6)$$

where λ_1 and λ_2 are two positive regularization coefficients. In general, the optimization in (6) is nonconvex due to the nonlinear coupling term $U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}, \mathbf{A})$. Thus, in the sequel, we exploit successive convex approximation (SCA) methods [26] to devise an iterative algorithm that converges to local solutions of (6). The main steps of the method are shown in Algorithm 1. The first step is an SCA optimization, where the nonconvex term $U_{\mathcal{X}, \mathbf{A}}(\cdot)$ in (6) is replaced by a convex approximation $\tilde{U}_{\mathcal{X}, \mathbf{A}}(\mathbf{x}, \mathbf{A}; \mathbf{x}(t), \mathbf{A}(t))$ evaluated at $(\mathbf{x}(t), \mathbf{A}(t))$. Among the possible choices for the approximant (see [26]), in this work we consider: $\tilde{U}_{\mathcal{X}, \mathbf{A}}(\mathbf{x}, \mathbf{A}; \mathbf{x}(t), \mathbf{A}(t)) = \nabla_{\mathbf{x}} U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}(t), \mathbf{A}(t))^T \mathbf{x} + U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}(t), \mathbf{A}(t))$. The problem in (7) admits a closed form solution: as for (5), $\mathbf{A}(t+1)$ is obtained by selecting the edges corresponding to the K smallest pairs among the potentials in $\lambda_1 \cdot U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}(t), \mathbf{A}(t)) + \lambda_2 \cdot U_{\Gamma}(\mathbf{A}, \gamma)$; whereas, we have $\hat{\mathbf{x}}(t) = \mathbf{y} + \lambda_1 \cdot \nabla_{\mathbf{x}} U_{\mathcal{X}, \mathbf{A}}(\mathbf{x}(t), \mathbf{A}(t))$. Finally, the second step in Algorithm 1 performs a convex combination between the previous signal estimate $\mathbf{x}(t)$ and the current solution $\hat{\mathbf{x}}(t)$, using a diminishing step-size $\alpha(t)$ ². Our method generalizes other approaches available in the literature [13], [15], [28], [29] which do not consider the community term U_{Γ} and assume only quadratic potentials for $U_{\mathcal{X}, \mathbf{A}}(\cdot)$.

IV. NUMERICAL SIMULATIONS

In this section, we assess the performance of procedures leveraging the proposed CMRF. Let us first consider learning

² $\alpha(t)$ satisfies $\sum_{t=0}^{\infty} \alpha(t) = \infty$ and $\sum_{t=0}^{\infty} \alpha(t)^2 < \infty$, which are well known conditions in the adaptive filters and control literature [27].

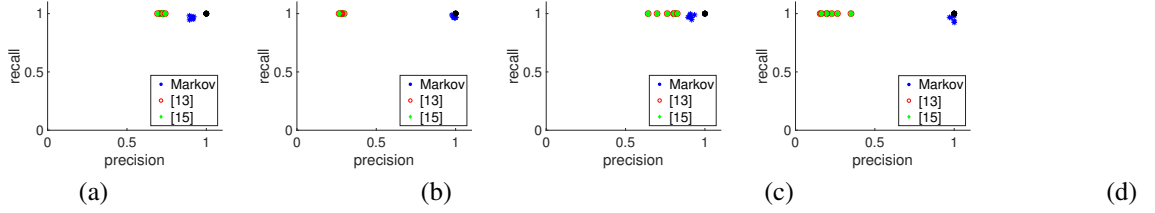


Fig. 2. Precision and recall for different values of signal variability ν (disassortative (a) and assortative (b) networks, $\nu = \{0, 0.5, 1, 1.5, 2, 2.5\}$) and for different number of communities P (disassortative (c) and assortative (d) networks, $P = \{3, 4, 5, 6, 7, 8, 9\}$). The black circle in (1, 1) represents perfect detection. The CMRF based method achieves uniformly better performances while being less sensitive to the parameter settings.

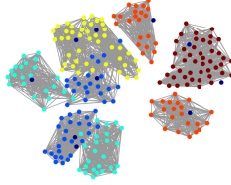


Fig. 3. Estimated edges over real social network data [30].

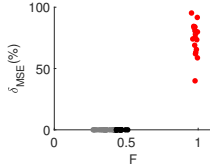


Fig. 4. Joint topology and signal recovery: percentage MSE reduction vs F-measure for Algo. 1 (red), [15] (black) and [13] (gray), 20 Montecarlo runs.

of random modular graphs with of $N = 30$ nodes and $P = 4$ communities, with intra/inter community link probability π_{intra}, π_{inter} . The signal in the i -th community is uniformly distributed in the interval $[i, i + \nu]$, $i = 1, \dots, P$. Both signal and community values are assumed to be known. In Fig.2 we report precision and recall³ of the proposed CMRF method, and other state-of-the-art learning methods [13] and [15], as well as the perfect detection point (1, 1) (black circle). Figs.2(a)-(b) correspond to different signal variability ν for (a) a disassortative ($\pi_{intra} = 0.01$) and (b) an assortative network ($\pi_{intra} = 0.99$)⁴; the number of networks edges differ, being larger in the dissortative case. Furthermore, in Figs.2(c)-(d) we illustrate precision and recall results for different number P of communities for disassortative (c) and assortative (d) network structure. As we can see from Fig. 2, the proposed CMRF outperforms the competitors in [13] and [15] for several values of the parameters, being systematically closer to the perfect detection point (1, 1). Also, Table I illustrates the F-measure, given by $F = 2 \cdot \frac{r \cdot p}{r + p}$ (the higher the better), averaged over 20 Montecarlo runs, for $P = 4$ and $\nu = 3$. The results in Table 1 show how the proposed method based on CMRF outperforms competitors on both assortative and disassortative networks.

As a further example, we report the performance of the

³Precision is the number of true detected edges over the total detected; recall is the number of true detected edges over the number of true edges.

⁴For both the networks $\pi_{intra} = 1 - \pi_{inter}$.

TABLE I
F-MEASURE (20 MONTECARLO RUNS, $N = 30, P = 4, \nu = 3$).

F-measure	CMRF	[15]	[13]
Assortative	0.98	0.46	0.36
Dissortative	0.93	0.87	0.5

algorithm on a real social network [30], composed of $N = 236$ nodes representing the children of a primary school; the true edge weights are associated to daily children interaction, the signal values and community labels are given by the children ages and attended classes. The CMRF based graph learning achieves precision $p = 0.76$, recall $r = 0.60$, and an overall F-metric $F = 0.68$. The SoG and the estimated links are illustrated in Fig.3⁵ For the sake of comparison, the algorithm in [13] achieves $F = 0.2$ on the same data.

Finally, we test the performance of Algorithm 1 applied to a random modular graph with $N = 30$ and $P = 4$. The signal \mathbf{x} in the i -th community is set to i and corrupted with white Gaussian noise of variance 0.5. We consider the Mean Squared Error (MSE) between the true value \mathbf{x} and the estimated signal $\mathbf{x}(t)$, i.e., $MSE(t) = \|\mathbf{x} - \mathbf{x}(t)\|^2$, and its reduction $\delta_{MSE} = (MSE(0) - MSE(t_{MAX})) / MSE(0)$ achieved at convergence by Algorithm 1 (where $t_{MAX} = 10$). In Fig.4 we report the scatterplot of δ_{MSE} obtained in 20 Montecarlo simulations, versus the corresponding F-measure (red points). As we can see from Fig. 4, Algorithm 1 achieves an F-measure close to one; the MSE reduction is above 50% in all runs. For the sake of comparison, in Fig.4 we also report the F-measure achieved by the methods in [15] (black points) and [13] (gray points). Since these approaches do not perform signal recovery, their δ_{MSE} has been set to 0. Clearly, the holistic approach of the proposed CMRF-based method leads to better performance in graph topology recovery even in presence of noisy signals, where conventional signal-based approaches are not robust.

V. CONCLUSIONS

In this paper, we have proposed a novel compound MRF (CMRF) for signals on graphs, jointly modeling signal values, edge weights, and community labels. The CMRF generalizes Markovian SoG models in the literature, addressing different graph structures and signal smoothness priors. We have applied the CMRF to two GSP inference tasks: graph learning, and joint topology and signal recovery. Numerical simulations on synthetic and real data illustrate the superior performance of our method with respect to state-of-the-art competitors.

⁵Different communities are represented by different node colors, apart from communities built by one node that are assigned a dark blue color.

REFERENCES

- [1] D.I. Shuman, Sunil K. Narang, P. Frossard, A. Ortega, P. Vandergheynst. "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains." *IEEE signal processing magazine* 30, no. 3 (2013): 83-98.
- [2] A. Ortega, P. Frossard, J. Kovacevic, J.M.F. Moura, P. Vandergheynst. "Graph signal processing: Overview, challenges, and applications." *Proceedings of the IEEE* 106, no. 5 (2018): 808-828.
- [3] M. Tsitsvero, S. Barbarossa, P. Di Lorenzo. "Signals on graphs: Uncertainty principle and sampling." *IEEE Transactions on Signal Processing* 64.18 (2016): 4845-4860.
- [4] G. Mateos, S. Segarra, A.G. Marques, A. Ribeiro. "Connecting the dots: Identifying network structure via graph signal processing." *IEEE Signal Processing Magazine* 36, no. 3 (2019): 16-43.
- [5] G. Luo, Y. Zhu B. Guo, "Fast MRF-Based Hole Filling for View Synthesis," in *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 75-79, Jan. 2018.
- [6] X. Meng, Z. Zhang, Z. Geng, M. Zhang, "MRF-Based Disparity Upsampling Using Stereo Confidence Evaluations," in *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 561-565, April 2018.
- [7] A. Habbouchi, M.E.Y. Boudaren, A. Aa issani, W. Pieczynski, "Unsupervised Segmentation of Markov Random Fields Corrupted by Nonstationary Noise," in *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1607-1611, Nov. 2016
- [8] W. Jamal, S. Das, I. Oprescu, K. Maharatna, "Prediction of Synchronstate Transitions in EEG Signals Using Markov Chain Models," in *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 149-152, Feb. 2015.
- [9] M.K. Reddy, K.S. Rao, "Robust Pitch Extraction Method for the HMM-Based Speech Synthesis System," in *IEEE Signal Processing Letters*, vol. 24, no. 8, pp. 1133-1137, Aug. 2017.
- [10] R. Warnick, M. Guindani, E.Erhardt, E. Allen, V. Calhoun, M. Vanucci, "A bayesian approach for estimating dynamic functional network connectivity in fmri data." *Journal of the American Statistical Association* 113.521 (2018): 134-151.
- [11] B. Nettasinghe, V. Krishnamurthy, "Influence Maximization Over Markovian Graphs: A Stochastic Optimization Approach," in *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 1, pp. 1-14, March 2019.
- [12] S. Jafarizadeh, "Fastest Mixing Reversible Markov Chain: Clique Lifted Graphs & Subgraphs," in *IEEE Transactions on Signal and Information Processing over Networks*.(Early Access)
- [13] H. E. Egilmez, E. Pavez, A. Ortega, "Graph learning from data under Laplacian and structural constraints." *IEEE Journal of Selected Topics in Signal Processing* 11, no. 6 (2017): 825-841.
- [14] S. Colonnese, M. Biagi, R. Cusani, G. Scarano, "Graph Adjacency Matrix Learning for Irregularly Sampled Markovian Natural Images", 2017 25th European Sig. Proc. Conf. (EUSIPCO)
- [15] X. Dong, D. Thanou, P. Frossard, P. Vandergheynst, "Learning laplacian matrix in smooth graph signal representations", *IEEE Tr. on Signal Processing*, 64(23), 2016.
- [16] B. Girault, S.N. Shrikanth, A. Ortega. "Local stationarity of graph signals: insights and experiments." In *Wavelets and Sparsity XVII*, vol. 10394, International Society for Optics and Photonics, 2017.
- [17] S. Colonnese, G. Pagliari, M. Biagi, R. Cusani, G. Scarano. "Compound Markov Random Field Model of Signals on Graph: An Application to Graph Learning." In 2018 7th European Workshop on Visual Information Processing (EUVIP), pp. 1-5. IEEE, 2018.
- [18] S. Geman, D. Geman. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". *IEEE Tr. on Pattern Anal. Mach. Intell.* 6, 6 (November 1984),
- [19] P. Brémaud, "Discrete Probability Models and Methods: Probability on Graphs and Trees, Markov Chains and Random Fields, Entropy and Coding", Springer International Publishing 2017
- [20] S. Colonnese, S. Rinauro, G. Scarano, "Bayesian image interpolation using Markov random fields driven by visually relevant image features," *Signal Processing: Image Communication*, Volume 28, Issue 8, 2013.
- [21] J. Besag "Spatial interaction and the statistical analysis of lattice systems.", *Journal of the Royal Statistical Society: Series B (Methodological)*, 1974 Jan;36(2).
- [22] G. Ortiz-Jimenez, M. Coutino, S.P. Chepuri, G. Leus "Sampling and Reconstruction of signals on product graphs." In 2018 IEEE Global Conf. on Signal and Information Proc. (GlobalSIP), pp. 713-717. IEEE, 2018.
- [23] A. Germain, J. Chakareski "Spotify me: Facebook-assisted automatic playlist generation", IEEE 15th International Workshop on Multimedia Signal Processing (MMSP) 2013.
- [24] J.O. Garcia, A. Ashourvan, S. Muldoon, J.M. Vettel, D.S. Bassett. "Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function." *Proc. of the IEEE* 106, no. 5 (2018): 846-867.
- [25] R.A. Horn, C.R. Johnson, "Matrix Analysis", *Cambridge University Press*, 2012
- [26] G. Scutari, F. Facchinei, L. Lampariello, "Parallel and distributed methods for constrained nonconvex optimization Part I: Theory." *IEEE Transactions on Signal Processing*, vo. 65, no. 8, pp. 1929-1944, 2016.
- [27] Ali H. Sayed, *Adaptive filters*, John Wiley & Sons, 2011.
- [28] S.P. Chepuri, Sijia Liu, G. Leus, A.O. Hero. "Learning sparse graphs under smoothness prior." In 2017 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6508-6512. IEEE, 2017.
- [29] P. Berger, M. Buchacher, G. Hannak and G. Matz, "Graph Learning Based on Total Variation Minimization," 2018 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, 2018, pp. 6309-6313.
- [30] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J. Pinton, M. Quagiotto, W. Van Den Broeck, C. Regis, B. Lina, and P. Vanhems, "High-resolution measurements of face-to-face contact patterns in a primary school." *PLoS one*, vol. 6, no. 8, p. e23176, 2011.