Abstract—Finite impulse response (FIR) graph filters play a crucial role in the field of signal processing on graphs. However, when the graph signal is time-varying, the state of the art FIR graph filters do not capture the time variations of the input signal. In this work, we propose an extension of FIR graph filters to capture also the signal variations over time. By considering also the past values of the graph signal, the proposed FIR graph filter extends naturally to a 2-dimensional filter, capturing jointly the signal variations over the graph and time. As a particular case of interest we focus on 2-dimensional separable graph-temporal filters, which can be achieved at the price of higher communication costs. This allows us to give the filter specification and perform the design independently in the graph and temporal domain. The work is concluded by analyzing the proposed approach for stochastic graph signals, where the first and second order moments of the output signal are characterized.

I. INTRODUCTION

Signal processing on graphs emerged recently as a tool to extend classical signal processing concepts from time and image signals to signals that reside on the vertices of an irregular graph. The breakthrough in this area was the definition of the graph Fourier transform (GFT) [1]–[3], which extended the graph signal analysis in the graph frequency domain. By having a specific definition of graph frequency, graph filters emerged as a basic building block. In this context, a graph filter is the direct analog of a classical temporal filter, now operating on the graph signal by amplifying or attenuating part of its graph spectrum. Graph filters have been used in applications like data classification and customer behavior prediction [1], signal denoising and smoothing [4], [5], solving consensus problems [6], anomaly and event boundary detection [7], [8], to name a few. Finite impulse response (FIR) graph filters appeared first with the property of having a polynomial frequency response, and thus can be easily implemented in the node domain [1], [9]. Secondly, infinite impulse response (IIR) graph filters were proposed due to the necessity to achieve better interpolation or extrapolation properties around the known graph frequencies [8], [10] and [11] dealing with the stability issues. However, the aforementioned works focus mainly on time-invariant graph signals, whereas time can carry extra information, for example financial time series of companies or goods in a stock market, temperature measurements taken continuously in time by a sensor network, or political popularity in social media. For these cases we can be interested to compute predictions, statistics or make inferences on this signal. This can potentially be more accurate when the time dependency of the signals is taken into account. Such aspects are catching attention recently, also in the signal processing on graph area [11], [12].

In [11], an autoregressive moving average (ARMA) recursion is analyzed to implement IIR graph filters. In case of a time-varying graph signal, the ARMA filter extends to a 2-dimensional filter and captures jointly the graph and temporal variations of the signal. However, designing stable 2-dimensional ARMA filters seems challenging. To avoid the latter problem related to IIR filters, we present an extension of FIR graph filters to capture also the time-variations of the graph signal. By incorporating also a temporal memory while computing the filter output, the well known FIR graph filters extend to 2-dimensional filters operating jointly on the graph and temporal spectral domain. In contrast to [11] the proposed 2-dimensional filter has more degrees of freedom to approximate a desired frequency response. Further, as a particular case we analyze the class of separable 2-dimensional frequency responses. This property allows us to give filter specifications and perform the design independently in the graph and temporal domain.

The letter is concluded by analyzing the proposed approach for a time-varying random process over the graph. For this case we calculate in closed form the first and second order moments of the filter output and we show that, in the mean, the proposed filter behaves as the same filter operating on a deterministic signal being the mean of the graph process.

II. PRELIMINARIES

Let us consider an undirected and connected graph $G$ of $N$ nodes. We indicate with $x \in \mathbb{R}^N$ the graph signal and with $L$ the graph Laplacian [3]. The GFT $\hat{x}$ of $x$ and its inverse are calculated as

$$\hat{x}_n = \langle x, \phi_n \rangle,$$

$$x_n = \langle \hat{x}, \phi_n \rangle,$$

respectively, where $\langle , \rangle$ denotes the inner product and $\{\phi_n\}_{n=1}^N$ are the Laplacian’s eigenvectors. The corresponding eigenvalues $\{\lambda_n\}_{n=1}^N$ form the graph frequencies. To avoid any restrictions on the applicability of the proposed approach, we present the results for a general Laplacian matrix $L$. We only require $L$ to be symmetric and local: for all $i \neq j$, $L_{ij} = 0$ whenever the nodes $u_i$ and $u_j$ are not neighbours and $L_{ij} = L_{ji}$ otherwise. We redirect the reader to [1]–[3] for more details.1

A graph filter $H$ is defined as a linear operator that acts on a graph signal $x$ by amplifying or attenuating different parts of its spectrum as

$$Hx = \sum_{n=1}^N H(\lambda_n)\langle x, \phi_n \rangle \phi_n.$$  

Let $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$ be the minimum and the maximum eigenvalues of $L$. The graph frequency response $H$:

1Note that the core idea can be applied also to directed graphs using the adjacency matrix instead of the Laplacian.
$[\lambda_{\text{min}}, \lambda_{\text{max}}] \rightarrow \mathbb{R}$ controls how much $H$ amplifies the signal component of each graph frequency

$$H(\lambda_n) = \langle Hx, \phi_n \rangle / \langle x, \phi_n \rangle.$$  

Given a desired graph frequency response $H^*(\lambda)$, the filter coefficients are found by solving a linear system when the graph and thus the frequencies $\lambda_n$ are known [1], [3]. In case the graph structure is not known, we can approximate $H$ by using a $K$-th order polynomial of $L$ and the filter output is

$$y = \left( a_0 \mathbf{I} + \sum_{k=1}^{K} a_k L^k \right) x.$$  

In this case, the filter coefficients are found for a continuous range of $\lambda$, they describe the graph frequency response of that filter for any graph [9]. The $\text{FIR}_K$ graph filters can also be computed distributively, since $L^K x = L(L^{K-1} x)$ and each node can compute the $K$-th term from the values of the $(K-1)$-th term in its neighbourhood. In this communication we will focus our attention to the universal case, where the filter frequency response is analyzed for a continuous range of graph frequencies $\lambda$ and we will address also the distributed implementation. However, notice that this is not with loss of generality, since a centralized implementation over a known graph topology is also possible.

### III. FIR GRAPH FILTERS WITH TIME-VARYING INPUT

In this section, we present our recursions that implement 2-dimensional graph-temporal filters. The first approach implements a 2-dimensional $\text{FIR}_K$ filter requiring $K$ times more data exchanges and computational complexity than the pure FIR graph filter (4). Then, we move to the more general approach, which requires $K^2$ more data exchanges and computational power to implement a 2-dimensional FIR filter, but it offers a more complete 2-dimensional transfer function with more degrees of freedom. Further, as a particular subclass of interest of the latter we propose 2-dimensional filters which have a separable frequency response in the graph and temporal domain.

#### 2D FIR graph filters.

Temporal variations of the input signal can be captured by the FIR filter taking into account its temporal history. Let us for instance consider the intuitive extension of (4) to this case:

$$y_t = \sum_{k=0}^{K} a_k L^k x_{t-k},$$

where now we can see that the output at time $t$, i.e., $y_t$, depends on the past $K$ realizations of the input signal, where $x_{t-k}$ is graph-shifted with $L^k$ (this favors a distributed implementation). Recursion (5) provides the intuition that it represents a $K$-th order FIR filter in both the graph and temporal frequency domain. To see this, we calculate the joint transfer function of the filter (5). Applying first the GFT and then the $z$-transform to (5), the joint graph-temporal transfer function can be written as

$$H(z, \lambda) = \sum_{k=0}^{K} a_k \lambda^k z^{-k}.$$  

We can now formally see that, the joint transfer function (6) implements an FIR filter of order $K$ in the graph domain, as well as, an FIR filter of the same order in the temporal domain. It can be noticed that it requires $K$ times more data exchanges and computational power than the classical $\text{FIR}_K$ graph filter. Being an FIR type of filter, it avoids the stability problems that a 2-dimensional ARMA filter presents [11]. However, from (6), we observe that the zeros of the polynomial in $\lambda$ and in $z$ are correlated to each other. This affects the joint design, and thus the approximation accuracy, where a tradeoff has to be found between the filter approximations in each domain.

We illustrate this in Fig. 1, where we approximate with an $\text{FIR}_3$ filter an ideal step function in the graph frequency domain with cut-off frequency $\lambda_c = 0.5$. We can see that for a high normalized temporal frequency the filter response differs from the case of $f = 0$, for which the filter has been designed. This behavior can be addressed to the fact that the joint transfer function (6) is not a complete 2-dimensional polynomial of order $K$. Indeed, all the cross term monomials of the form $\lambda^\alpha z^{-\beta}$ with $\alpha \neq \beta$ are missing. However, even considering all these challenges, such an approach can still be used to approximate some specific 2-dimensional filter masks. Similar considerations are valid also for the ARMA filters [11], which render both approaches suitable for graph signals that varies slowly in time ($f \approx 0$), where its graph frequency response is similar to the static case ($f = 0$).

The approximation accuracy of the 2-dimensional FIR filter can be improved if we incorporate also the missing cross term monomials in (5). This can be achieved by considering all $K$ graph-shifts for every past input of the graph signal. Thus, this approach considers more data exchanges and computational power to implement a filter of the same order in both the graph and temporal domain. More formally, consider the recursion

$$y_t = \sum_{k=0}^{K} \sum_{l=0}^{K} a_{k,l} L^k x_{t-l},$$

with $K_g$ and $K_t$ the memory of the filter in the graph and temporal domain, respectively. We can now calculate the joint transfer function of (7) in the same way as we did for (5). By applying the GFT and the $z$-transform we obtain

$$H(z, \lambda) = \sum_{k=0}^{K} \sum_{l=0}^{K} a_{k,l} \lambda^k z^{-l},$$

which is now an FIR of order $K_g$ in the graph domain and of order $K_t$, not necessarily equal to $K_g$, in the temporal domain.

In case of ARMA filters this is observed for the poles, which affects both the filter design and stability.
From the joint frequency response (8) we can see that now we have a full polynomial in the variables $\lambda$ and $z$. Thus with this expression we can operate on all the $K_K K_t$ coefficients $a_{k,t}$, instead of the $K$ offered by (5), to approximate a given two-dimensional frequency response. Further, with respect to (5) and to the ARMA filter [11], recursion (7) has the potential to achieve filters of different orders in each domain. It can be noticed, that to implement the 2-dimensional FIR$_K$ recursion (7), $K^2$ more data exchanges between nodes and computational power are required. To be implemented in a distributed way, the nodes must also exchange all their information before the graph signal changes. Meanwhile, this is not required in a centralized approach, where all the data is available. Given the general form (7) and its particular version (5) there is room to use an intermediary approach, which can potentially improve the approximation accuracy of (5) still preserving the same distributed implementation cost. The best (i.e., with more degrees of freedom) 2-dimensional distributed filter that we can realize is

$$y_t = \sum_{k=0}^{K_t} \sum_{l=0}^{K_t} a_{k,l} L^k x_{t-l},$$

which is similar to the causal autoregressive model presented in [12], yet without any stability issue. In the rest of this communications, we will address a particular subclass of interest of (7) which has the property to achieve a separable 2-dimensional frequency responses in graph and time. By setting $a_{k,t} = b_k c_l$ we have

$$y_t = \sum_{k=0}^{K_t} \sum_{l=0}^{K_t} b_k c_l L^k x_{t-l},$$

where $b_k$ are the filter coefficients relative to the graph part and $c_l$ to time. Similar to the derivation of (7), the transfer function of (10) can then be written as $H(z, \lambda) = H_t(z) H_g(\lambda)$ where $H_t(z) = \sum_{k=0}^{K_t} c_k z^{-l}$ and $H_g(\lambda) = \sum_{k=0}^{K_t} b_k \lambda^k$. Note that even having a separable 2-dimensional time-graph frequency response, (10) operates jointly on the graph and time domain. This separable approach, as we will see later, offers us the freedom to handle the filter specifications independently in the graph and temporal domain. On the other hand, the separable filters are limited w.r.t. the general case (7) since we now have $K_t + K_t$ degrees of freedom. Thus it can address a limited class of 2-dimensional frequency responses, but a very practical class.

Filter design problem. Without loss of generality, we present the design problem of a 2-dimensional FIR graph-temporal filter in order to approximate a desired frequency response $H^\ast(e^{j\omega}, \lambda)$ over a continuous range of frequencies in both the graph and the temporal frequency domain. This design approach, like the one performed only on the graph domain [9], is useful when the graph spectrum is unknown, when computing the eigenvalue decomposition of the Laplacian has an unaffordable cost, or when we are interested in finding the filter coefficients independently from the structure of the underlying graph. Consider that a desired 2-dimensional frequency mask $H^\ast(e^{j\omega}, \lambda)$ is given and we are interested in finding the filter coefficients in order to approximate it. For the non separable cases, this can be achieved by a 2-dimensional polynomial fitting of $H(z, \lambda)$ in (6) or (8) (for $z = e^{j\omega}$) to $H^\ast(e^{j\omega}, \lambda)$.

For the case that the desired frequency response is separable, i.e., $H^\ast(e^{j\omega}, \lambda) = H^\ast_t(e^{j\omega}) H^\ast(\lambda)$ we have the benefit to separate the filter design as well. Thus, we can use any desired method to find the coefficients $b_k$ that approximate $H^\ast(\lambda)$, as well as any of the well-known techniques to find the coefficients $c_l$ for approximating $H^\ast_t(e^{j\omega})$. The latter, renders the separable approach very practical, since we can give our specifications independently in the graph and temporal domain. Further, it ensures a approximation accuracy of the filter specifications.

IV. STOCHASTIC ANALYSIS

We now analyze the FIR filter behavior when the graph signal has a stochastic nature over time. This may happen, for instance, when the signal on the graph is corrupted by noise. Similar to the ARMA graph filter [13], we characterize statistically the 2-dimensional FIR when the graph signals has a temporally non-stationary mean but a temporally stationary covariance. For our analysis we consider the following signal model.

Random signal model. The graph signal $x_t$ at time instant $t$ is a realization of a random process with time-varying first order moment $\bar{x}_t$ and time-invariant covariance matrix $\Sigma_{xx}$. $\bar{x}_t$ is independent over time.

The above random signal model generalizes the deterministic signals analyzed in current literature, and basically tells us that the graph signal might be correlated among the nodes for a fixed time instant $t$, but has independent realizations with different means over time. It can be seen, for instance, as a desired time-varying signal $\bar{x}_t$ embedded in noise, more specifically in the form $x_t = \bar{x}_t + n_t$, with $n_t$ being zero mean noise with covariance matrix $\Sigma_{xx}$. With this in place, the following can be claimed.

Proposition 1: Consider a separable 2-dimensional FIR filter of orders $K_t$ and $K_t$ in the graph and temporal domain, respectively, and consider a graph signal that follows the proposed signal model. Then, the expected value $\bar{y}_t$ and the covariance matrix $\Sigma_{yy}$ of the output signal are given by

$$\bar{y}_t = \sum_{k=0}^{K_t} \sum_{l=0}^{K_t} b_k c_l L^k \bar{x}_{t-l}$$

and

$$\Sigma_{yy} = ||c||^2 \sum_{k=0}^{K_t} \sum_{m=0}^{K_t} b_k b_m L^k \Sigma_{xx} (L^m)^\top,$$

where $|| \cdot ||$ indicates the 2-norm and $c = [c_1, c_2, \ldots, c_{K_t}]^\top$.

Proof: (Sketch) The results can be proven by applying the definitions of the expected value and the covariance matrix to the separable version of (7).

Proposition 1 extends the analysis of 2-dimensional FIR filters to a stochastic environment. It says that, in the mean, the FIR filter behaves as the same 2-dimensional filter operating on a deterministic time-varying signal, being the time-varying

3This aspect will be covered in more detail in future research.

4We derive the statistics for the separable case of interest, but the same can be derived also for (7).
mean of the input graph process. Further, it gives us a closed-form formula to calculate the covariance matrix of the output signal in order to see how far from the mean a given realization can be. We can notice from (11b) that the variance of the output signal, at each node, depends on the squared norm of the temporal filter taps. This gives us a handle on the variance of the output signal by tuning these coefficients. In case of a stationary process, the result collapses to the 1-dimensional FIR filter operating on the mean signal, like in [13].

V. NUMERICAL EVALUATION

To illustrate our conclusions, we start by showing that recursion (10) can approximate different separable filters with given specifications in the graph and temporal frequency domain. Then, we use the 2-dimensional FIR filter to denoise a time-varying signal which is also affected by interference. Finally, we simulate a scenario where the graph signal is a stochastic process with a time-varying mean. For the filter design phase, we use the polynomial approximation [9] for the graph domain and the windowing method for the time domain [14]. The results are derived for a 2-dimensional FIR filter of orders $K_g = K_t = 10$. The results are carried out over a graph of 100 nodes randomly placed in a squared area, with two nodes being neighbors if they are physically closer than 15% of the maximum distance in the area.

Filter approximation. With reference to Fig. 2, we can see that the proposed approach can approximate different desired 2-dimensional separable frequency responses. For this particular case, the cut-off frequencies in both domains are chosen as the half of the respective bands.

Denoising and interference cancellation. Consider a graph signal of the form $x_t = s_t + i_t + n_t$, where

$$\langle s_t, \phi_t \rangle = \left\{ \begin{array}{ll} e^{j \pi t / 4} & \text{if } \lambda_n < 0.5 \\ 0 & \text{otherwise} \end{array} \right. \quad (12)$$

is the signal of interest, $\langle i_t, \phi_t \rangle = e^{j \pi t / 4}$, $\forall \lambda_n$, is the interfering signal and $n_t$ is a zero mean additive Gaussian noise with $\Sigma_{xx} = \sigma^2 I$ and $\sigma^2 = 0.1$. Our goal is to recover the graph signal of interest $s_t$ using the 2-dimensional FIR graph filtering approach. In this way we aim to use the FIR filter to cancel the out of band noise in the graph domain and cancel the interferer in the temporal domain.

To measure the performance of our solution, we define the following two errors

$$\epsilon_{t_{\text{interf}}} = \frac{\| \hat{y}_t - \hat{y}_{t}^* \|}{\| \hat{y}_{t}^* \|}, \quad \epsilon_{t_{\text{total}}} = \frac{\| y_t - \hat{s}_t \|}{\| \hat{s}_t \|}. \quad (13)$$

where $\hat{y}_t$ and $\hat{y}_{t}^*$ are the graph Fourier transforms of the filter outputs at time $t$ when the input signals are $x_t$ and $s_t + n_t$, respectively. The first error, $\epsilon_{t_{\text{interf}}}$, is a measure on how good we attenuate the interfering signal. Indeed, it tells us how well we approximate the filter output $\hat{y}_t^*$ when there is no interference. Meanwhile, the second error, $\epsilon_{t_{\text{total}}}$, tells us how good we suppress the interference and the noise. For our simulations, the coefficients $b_k$ are designed to approximate an ideal low pass step function with $\lambda_c = 0.5$, meanwhile the coefficients $c_l$ are found by approximating a low-pass temporal filter with cut-off frequency $\omega_c = \pi/2$. In Fig. 3, we can see that after some initial assessment time of the filter, both errors reduce. Specifically $\epsilon_{t_{\text{interf}}}$ is reduced by an order of two.$^5$

This shows the robustness of the 2-dimensional FIR filter (7) to interference, where the interfering signal is attenuated by the filter. An FIR graph filters that takes the input signal only once, in the beginning of the filtering, produces errors which are much higher due to their impossibility to operate on the temporal frequencies.

Stochastic analysis. We now aim to illustrate the theoretical results obtained in Section IV. With the same setup as in the previous paragraph, we plot in Fig. 4 the output signal of the filter and its analytical expected value (11a) as a function of time for two representative nodes. We can notice that the output signal oscillates with the same frequency of the desired input signal around its expected value. Further, due to the fact that we attenuate the noise out of the band of interest we have also a reduction of the signal variability around its mean. The theoretical/empirical variance for node 5 is 0.0074/0.0069 and for node 100 is 0.0140/0.0149. These results show also that we attenuate the noise out of the band of interest we have a reduction of the signal variability around its mean. The empirical variance is calculated over 1000 samples.

Future research will analyze the proposed approach on time-varying graphs and testing it on real scenarios.

$^5$Notice that the filter can also be designed as a band-pass around the oscillating frequency of $s_t$. This will further help to reduce the noise in the temporal frequency domain.
REFERENCES

[1] A. Sandryhaila and J. M. Moura, “Discrete Signal Processing on Graphs,” IEEE Transactions on Signal Processing, vol. 61, no. 7, pp. 1644–1656, 2013.
[2] ——, “Discrete Signal Processing on Graphs: Frequency Analysis,” IEEE Transactions on Signal Processing, vol. 62, no. 12, pp. 3042–3054, 2014.
[3] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and 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, vol. 30, no. 3, pp. 83–98, 2013.
[4] S. Chen, A. Sandryhaila, J. M. Moura, and J. Kovacevic, “Signal Denoising on Graphs via Graph Filtering,” in IEEE Global Conference on Signal and Information Processing (GlobalSIP), Atlanta, GA, USA, December 2014, pp. 872–876.
[5] F. Zhang and E. R. Hancock, “Graph Spectral Image Smoothing Using the Heat Kernel,” Pattern Recognition, vol. 41, no. 11, pp. 3328–3342, 2008.
[6] A. Sandryhaila, S. Kar, and J. M. Moura, “Finite-Time Distributed Consensus Through Graph Filters,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 1080–1084.
[7] H. E. Egilmez and A. Ortega, “Spectral anomaly detection using graph-based filtering for wireless sensor networks,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 1085–1089.
[8] A. Loukas, M. A. Zúñiga, I. Protonotarios, and J. Gao, “How to Identify Global Trends from Local Decisions? Event Region Detection on Mobile Networks,” in International Conference on Computer Communications (INFOCOM), 2014.
[9] D. I. Shuman, P. Vandergheynst, and P. Frossard, “Chebyshev Polynomial Approximation for Distributed Signal Processing,” in IEEE International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011, pp. 1–8.
[10] X. Shi, H. Feng, M. Zhai, T. Yang, and B. Hu, “Infinite Impulse Response Graph Filters in Wireless Sensor Networks,” IEEE Signal Processing Letters, Jan 2015.
[11] A. Loukas, A. Simonetto, and G. Leus, “Distributed Autoregressive Moving Average Graph Filters,” Signal Processing Letters, vol. 22, no. 11, pp. 1931–1935, 2015.
[12] J. Mei and J. M. Moura, “Signal processing on graphs: Estimating the structure of a graph,” in Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on. IEEE, 2015, pp. 5495–5499.
[13] E. Isufi, A. Simonetto, A. Loukas, and G. Leus, “Stochastic Graph Filtering on Time-Varying Graphs,” in IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015.
[14] Digital Signal Processing Committee of the IEEE Acoustics, Speech, and Signal Processing Society, eds., Programs for Digital Signal Processing. IEEE Press, 1979, Algorithm 5.2.