Engineering Mathematics
2025; 9(1): 1-15
http://www.sciencepublishinggroup.com/j/engmath
doi: 10.11648/j.engmath.20250901.11
ISSN: 2640-0855 (Print); ISSN: 2640-088X (Online)

A Matrix Model for Adaptive Graph Filtering Using a Generalized Mean-sets Theory's Approach

Christophe Fotso[1,*], Duplex Elvis Houpa Danga[1], Daniel Tieudjo[2]

[1]Department of Mathematics and Computer Science, The University of Ngaoundere, Ngaoundere, Cameroon
[2]Department of Mathematics and Computer Science, National School of Agro-Industrial Science, Ngaoundere, Cameroon

Email address:
fotch2016@gmail.com (Christophe Fotso), elvishoupa@gmail.com (Duplex Elvis Houpa Danga), tieudjo@gmail.com (Daniel Tieudjo)
[*]Corresponding author

To cite this article:
Christophe Fotso, Duplex Elvis Houpa Danga, Daniel Tieudjo. (2025). A Matrix Model for Adaptive Graph Filtering Using a Generalized Mean-sets Theory's Approach. *Engineering Mathematics, 9*(1), 1-15. https://doi.org/10.11648/j.engmath.20250901.11

Received: 27 March 2025; **Accepted:** 8 April 2025; **Published:** 6 May 2025

Abstract: Integrating Mean-sets theory employing generalized graph or group-theoretic tools and techniques into adaptive graph filtering can lead to more effective resilient filtering processes, particularly in challenging environments with clutter or uncertainty. In this paper, we show that under some crucial smoothing assumptions, the generalized Mean-sets theory developped for negatively curved convex combination Polish metric spaces following the formal Means-sets probability theory's approach from Natalia Mosina, provides a new useful system for some secure adaptive graph filtering processes. We use convex combination operations (in the sense of Terán and Molchanov) on both individual input graph signals and filters. Individual adaptive graph filters being independently adapted by space (dataset)-valued random variables, while the convexification operator on the underlying dataset acts as a flexible theoretical instrument for preserving some good features of the standard scheme, like privacy of their informative trends, and looks more robust to changes. We exhibit a graph matrix model from a system of convex combination of two adaptive finite impulse response (FIR) graph filters processing a sampled-weighted mean-set (expectation) of some transversal graph signals with finite length $N \geq 2$.

Keywords: Graph Signal, Convex Combination, Mean-set, Convexification Operator, Secure Adaptive Filtering

1. Introduction

Graph filters are systems or information processing architectures that process signals on graphs. These computational devices attempt to model the relationship between two or more graphs signals in real time but in an iterative manner. In their processing, they almost always preserved as much as possible the relevant content of known labels commonly called input signals to predict unknown labels called output signals, for the task at hand in domains involving digital signals, where filtering is used to extract relevant patterns from the data. This topic is well adressed in [11]. Meanwhile lectures on tracking, clustering and filtering in the conventional signal processing setting can be found in [15] who adressed on tracking filters in heavily cluttered environments with low target detection probabilities. [2, 3] studied tracking performances of a convex combination of two adaptive least mean-square (LMS) filters. For the same type of filters, [6] dealt with adaptive filtering while [27] investigated secure adaptive filtering.

The literature reveals that the first attempts to use mean-sets theory arguments for adaptive filtering and Signal processing are attributed to the Indian Engeneer called Choudhury D. M. R., who established a mathematical framework for mean-sets that emphasizes central tendencies and their statistical properties and adaptability in continuous data. His researches focussed on developing the mathematical concept of mean-sets and exploring its utility in various scientific and engineering domains including adaptive filtering and signal processing. Mean-sets theory is shown to improve the efficiency, effectiveness and robustness in adaptive filtering, particularly in clutter environments with noise, variability and uncertainties. But nowadays, clustering and filtering of graph signals while preserving input data's privacy are still some fundamental problematics in both Graph data analysis and conventional Signal theory. And, they have been studied extensively in the literature, even if clustering and filtering theories that justify the use of several related algorithm models proposed are still unsatisfactory.

In 2009, Mosina developped a new approach she also called "Mean-sets probability theory" (or "Mean-sets theory" to make short). This represented a significant contribution to the construction of a solid mathematical basis in the setting of Probability theory in general and on abstract structures, such as graphs for finitely generated groups [16]. Also, the features of this new theory include comparative framework, as it offers a novel approach to comparing different probabilistic models based on their mean-sets représentations. Indeed, she developped her theory on locally finite graphs, and hence (via Cayley graphs) on a finitely generated groups. Thas is, she employed graph-theoretic and group-theoretic techniques to analyse structures, relationships and connectivity between graph elements. She focussed on combinatorial or discrete structures and relationships, specially in graphs and groups, within the context of algebraic and geometric properties, making this approach more suited for graph signal processing, network analysis and combinatorial optimization.

In 2024, Fotso *et al.* [7] proposed a generalization of Mosina's probabilistic approach to some general metric spaces, notably negatively curved convex combination Polish (NCCCP) metric spaces. In their work, they emphazized on statistical properties of sets of outcomes valued in graphs rather than individual probablities. This allowed a more flexible interpretation of probabilistic models, accommodating scenarios where conventional methods may fall short, such as in high-dimensional contexts (e.g., the fields like machine learning) or when dealing with uncertainties in data. This advanced framework extended traditional probability theory on combinatorial by emphasizing the idea of "mean-sets" in linear algebraic and geometric graphs structures. This innovative approach on metric spaces provided a more nuanced understanding of uncertainty and randomness, especially in complex systems.

In the framework of metric spaces like in machine learning, the development of clustering or filtering classic theory needs specific conceptual works when adressing some fundamental or technical questions. For general metric spaces particularly, such questions among others concern: "What is a cluster in a set of data points?" or "How do we find clusters and good partitions of clusters under such a definition?" Fortunately, Chang *et al.* recently attempted to tackle these two interesting questions in [5].

In this paper, we show how generalized Mean-sets theory's approach formally developped for NCCCP metric spaces can offer, under certain crucial conditions, a useful way to improve the graph filtering process and target adjustment. This approach based on k-means and convex combination (CC) methods (in the sense of Terán and Molchanov) on graphs provides a flexible instrument to explore the best reference graph's data from the spatial variables of the input and output data. We study the filtering property of an average (expectation) set, that is, a weighted empirical graph signal of a finite number of undividually secured signal of lenght $N \geq 2$, that performs close to the best of its component signals without degradation. Individual graph signals are adapted independently by spatially valued random variables on which the natural convexification operator acts to preserve the privacy of their informative tendencies. While the empirical weighted average objet obtained using a convexification function is adapted to minimize the error graph signal of the global structure.

The outline of this paper is organized in eight sections. Section 2 is some general overview on signal theory and our motivations. Section 3 presents the key contributions of our approach. Section 4 is a recall on Graph signal processing and some classical smoothing data association (SDA)'s method and technics. Here, we provide some preliminaries and general assumptions for a good matching to the extended SDA's technics. In Section 5, a complete description and some illustrations of our approach is provided. The main focused aspects being the privacy-preservation scenario en the grap filtering processing. In Section 6, an implementation of our approach is carried out. As result, a model is proposed and grounded in a specific theoretical framework, therefore provinding a mathematical tool for adaptive graph filtering process. In section 7, a theoretical study of an adaptive convex combination of two individual adaptive graph filters is carried out. The last Section is devoted to our conclusion.

2. General Overview and Motivations

Some topics involving k-means method, which operationally associates each observed data with the unique closest cluster center of the k clusters in an appropriate sense, tracking property or performance of adaptive filters, have been used by Arenas-García *et al.* in [2, 3]. Classical privacy problems of adaptive filtering has been tackled in many papers among which [27] looks more rich of informations and where some relevant solutions have been proposed, providing some guidelines for the privacy-preserving of signals that are processed by untrusted agents or in cluterred environments. On such environments, several authors have designed variants of convex combination methods for solving some nonlinear problems using different ways and means (see, e.g., [12, 15, 29]). For example, Arenas-García *et al.* showed that convex combination approaches can help to improve adaptive filters properties or performances [2, 3].

The complexity and the irregularity of such structured systems and interactions mean that datas now reside on complex and irregular graphs that do not lend themselves to standard signal processing tools. Therefore, these tools and concepts from the usual signal processing framework need to be extended to graph signals, i.e., signals whose samples are indexed by the nodes of arbitrary graphs. Let consider that a graph signal consists of (possibly weighted) vertices, where the data values are defined as objects residing on a set of nodes, and the edges are the links connecting these vertices or nodes [19]. One can find that graphs offer the ability to model such datas and their complex interactive links. Also, graphs exploit some fundamental links between datas based on their relevant characteristics. In graph signal processing, the first step is to define the graph as a signal domain. If data sensing points (graph vertices) are commonly well defined, that is not the case with their connectivity (graph edges).

Nowever, considering the use of convex combination methods for solving nonlinear problems in signal theory and extended graph signal theory like in interconnected environments, the following problematics arise: Is there a possible connexion between the generalized Mean-sets theory's approach for metric spaces according to Fotso et *al.* [7], which is essentially based on Terán and Molchanov's CC operation and the preserving privacy signals in cluttered environments? If it's the case, how can this approach help to solve some nonlinear problems in these domains?

To effectively tackle the privacy problems involving graph signal processing (GSP), an other research field called graph signal processing in the Encrypted domain has arised. And by the time, it has faced some challenges and applications like looking by a cryptographic point of view to efficiently apply privacy preserving technics to common signal processing operations. The theoretical grounds of signal processing in the Encrypted domain come from the field of secure function evaluations [27] and more general support generic function evaluations [22]. Remember that the existence of efficient solutions for a secure execution of some generic functions is still problematic. However, many efficient and secure technics have been developed for specific applications in the past years, building up a set of tools that foretell the potential of this technology [27].

In general graph theory, the traditional weighted graph consists of weights on edges only and in the classic setting of Graph signal, the irregularity of datasets makes them be view as signals represented by graphs, so that signal coefficients are indexed by graph nodes and relations between them are represented by weighted graph edges [13]. In this way, edge-weighted graphs can be entirely described by their so called adjacency matrix, denoted by $A = (w_{ij})$, where w_{ij} stands for the weight of the edge $i - j$ having as source vertex the node indexed by i and target vertex the node indexed by j. The cases of graphs that can be weighted on their vertices, e.g., finite simple graphs or complete graphs, can also be described by their so called vertex-adjacency matrix, having a different definition, we shall give later. It is evident that graphs offer the ability to model irregularly structured datasets and their complex interactive connections. Also, graphs exploit some fundamental links between datas based on their relevant characteristics. Note that in graph signal processing, the first step is to define the graph as a signal domain. While the data sensing points (i.e., graph vertices) are commonly well defined, that is not the case with their connectivity (i.e., graph edges). In our setting where weights are focused on vertices either than edges, a vertex-weighted graph signal is obtained, looking like a doubly-weighted graph.

Now, being awared of the use of CC methods for solving nonlinear problems in classic signal theory like in cluttered environments, the following problematics arise: Is there a possible connexion between the generalized Mean-sets theory's approach for metric spaces, which is essentially based on Terán and Molchanov's convex combination operation in graphs, and the graph filtering signals coupled with privacy problems in cluttered environments? If the case, how can this approach help to solve some nonlinear problems in these domains? A motivating element when looking for an answer to these questions is that, the original Mean-sets theory's arguments are based on the general concept of vertex-weighted graph, which is also of great importance in graph signal processing theory.

Moreover, the generalized mean-set approach developped in NCCCP metric spaces [7] provides a convexification function that combines the Terán and Molchanov's convexifiaction operator and energy conservation arguments. And, the generalized weighted mean-set (expectation), when it exists, is always a convex combination of some convexified datas, which then appears as limits for the law of large numbers. Also, the negative curvature property of the underlying space assigns a unique mean-set (expectation) to each space-valued random variable [7, 24, 26].

Another motivating element when looking for an answer to these questions is that, the formal Mosina's Mean-sets theory is based on the general concept of vertex-weighted graph, which is also of great importance in graph signal processing theory, where a signal (on a graph) is defined by associating real (or complex) values to each vertex at a given time. As in classical signal processing, graph signals can have properties like smoothness (appropriately defined). Graph signals can be weighted, filtered, sampled, averaged, or denoised. If the graphs cannot be directly observed, we can model them and learn their underlying structure from data. With GSP, one gains access to principled tools mimicking the classical ones [19]. Moreover, in NCCCP metric spaces (like in separable Banach spaces), the generalized mean-set approach developped in [7] provides a convexification function that can allow, after the Terán and Molchanov's convexification operator, to use energy conservation arguments. Moreover, the generalized weighted mean-set (expectation) when it exists is always a convex combination of some convexified datas, which then appears as limits for the law of large numbers. Also, the negative curvature property of the space in question assigns a unique mean-set (expectation) to each space-valued random variable [7, 24, 26].

3. Key Contributions of Our Approach

The novelties of our approach can be summarized in the following three aspects:

First, this method can help to average graph datas where the graph is directly observed as metrically weightable on vertices, or if not, its underlying structure can be learned from data in a NCCCP metric space. For this class of geometric structures, this approach is performed by using Terán and Molchanov's CC method to obtain the so called weighted mean-set (expectation) overall graph secured signal of a sample of mutually i.i.d. data signals indexed by nodes in a vertex-weighted graph, the same like weighted mean-set (expectation) for NCCCP metric spaces. More specifically, for a practical application, this study proposes a novel CC approach for vertex-weighted graph adaptive filtering while preserving data privacies or characteristic properties. This is a promising alternative to the existing CC approaches on usual edge-weighted graphs. With this, we study the adaptive graph filtering property of an empirical weighted mean-set (expectation) of N transversal individual adaptive filters ($N \geq 2$). The individual graph filters being independently shifted by the natural convexification operator of the underlying metric space acting for preserving privacy or normalizing their error signals, while the empirical weighted mean-set (expectation) obtained using convexification function is adapted to minimize the error signal of the overall graph structure.

Second, this study proposes a novel convex combination tool for efficient filtering while preserving data's privacy or characteristic properties. This is a promising alternative to the existing approaches in the fields of Signal and Graph signal processing. With this, we study the filtering property of an empirical weighted mean-set (expectation) of N transversal adaptive filters ($N \geq 2$). Individual filters being independently adapted by the natural convexification operator acting for preserving privacy or normalizing their own error signals, while the empirical weighted mean-set (expectation) obtained using convexification function is adapted to minimize the error signal of the overall structure. It is worth noticing that in our setting, the output overall secured adaptive graph filter is not simply a usual convex combination of indivduals input adaptive filters, but it is that of the convexifiers of these indivdual input secured graph signals. This means that, the privacy action on each transversal input graph signal has been executed before their fitting through a normalized convex combination operation (where the sum of weights/parameters equals to the total mass 1). Indeed, taking the Terán and Molchanov's convexification operateur $\mathbb{K}$ like a sort of secure function/operator defined on the objects space, one can more specifically see that, its algebraic and geometric properties of this operator allow to act on adaptive filters as objects while preserving certains characters.

Third, to the best of our knowledge, there is no study extending the generalized Mean-sets theory approach based on Terán and Molchanov's CC method for tracking, filtering and even fitting. Moreover, it is worth noticing that when analyzing the filtering properties of least mean-squares (LMS) filters it is customary to study the influence of the step-size on the performance of the filter [2]. However, in this paper we show that the generalized weighted mean-set(expectation) adaptive filter which is the adaptive convex combination LMS filters is able to improve over the fitting capabilities of its components filters.

4. Recalls on Graph Signal Processing and SDA's Methods

4.1. Graph Signals and Filtering

Extending concepts from classical signal processing to signal processing on graphs remains a great challenge, which has been addressed in many paper among which [11, 13, 19].

In general, filtering, clustering or fusioning while preserving input data's privacy are some fundamental problematics in Graph data analysis like in signal theory. And they have been studied extensively in the literature, even if clustering or filtering theories that justify the use of several related algorithms proposed are still unsatisfactory. One of the major difficulties in the application of conventional multitarget tracking changes involves the problem of associating target measurements with the appropriate tracks, especially when there are missing measurements, unknown clusters or high false alarm rates (clutter) [2]. But, for the case of singletarget tracking, one can't talk about data association as the problem just looks like fitting a curve to a set of points [25]. However, very often, some phenomena dealing with adaptive filter's tracking problems happen in a two-party setting where the first party concerns multitarget tracking with input of individual adaptive filters [6], and the second party is a singletarget tracking of their combination, i.e., the overall adaptive filter.

In the related art, some topics like k-means method, which associates each observed data with the unique closest cluster center of the k clusters in an appropriate sense, tracking properties/performances of adaptive filters have been studied by Arenas *et al.* [2, 3]. Classical privacy problem of adaptive filtering has been addressed by many papers [27], where some relevant solutions have been proposed, providing some guidelines for the privacy-preserving of signals that are processed by untrusted agents or in cluterred environments. On such environments, several authors have designed variants of CC methods for solving some nonlinear problems by using different ways and means [12, 15, 29]. For example, Arenas-García *et al.* [3] showed that convex combination approaches can help to improve adaptive LMS filters properties/performances.

Similarly to the conventional signal processing, the complexity and the irregularity of graph systems and interactions mean that datas now reside on complex and irregularly structured datasets, that do not lend themselves to standard signal processing tools. Let consider a graph signal consists of (possibly weighted) vertices, where data values are defined as objects residing on a set of nodes and the edges are the links connecting these nodes [19]. By graph filters we mean systems which take a graph signal as an input and produce another signal indexed by the same graph as the output.

Our arguments may be used for averaging datas using convex combination operation in some domain like Graph signal theory through extended SDA methods and technics. For the classic SDA topic, we recommand the reader to see, e.g., [2, 14, 15, 21, 25, 28] for some more details.

4.2. A Formal Introduction to Conventional SDA's Technics

We now present a brief mathematical description of SDA's technics for multitarget tracking which involved k-means methods and τ-convergence arguments.

Let $(X = \mathbb{R}^n, \|\cdot\|)$ be a separable space which is isomorphic to its topological bidual, with the norm $\|x - y\|^2 = \sum_{i=1,\dots n} (x_i - y_i)^2$. The problem of data association by smoothing is traduced in this context by the question: How can we do to smoothly associate spatial datas $\{(t_i, z_i)\}_{i=1}^{\ell} \subset [0,1] \times \mathbb{R}^n$ to unknown trajectories $\rho_j : [0,1] \longrightarrow \mathbb{R}^N$, for $j = 1, \dots, N$? In other words, what is the best strategy to distribute a data set $H_\ell = \{\xi_1, \dots, \xi_\ell\}$ of values taken by a $X = \mathbb{R}^N$-valued random variable $\xi : \Omega \longrightarrow X$, in a cluster of N targets, where each target has ρ_j as its center for $j = 1, \dots, N$?

Notice that, when $N = 1$, i.e., for the case of singletarget tracking, this problem involves fitting a curve to a set of points. Hereby, we are not yet talking about data association. But when $N > 1$, we are in the case of multitarget tracking. Therefore, both smoothing and data association problems are coupled.

In general, the centers of clusters may take their values in a space different from data's space. If the case, certain imperfections coming from tracking (true tracks and false tracks can exist), signaling or recording, and globally grouped within what is usually referred to as the ?noise phenomenon?, are to be controlled [25].

In our context, let Γ be a vertex-weighted metric graph where data points are indexed by nodes in space-time while cluster centers are functions from space-time to the vertex-set. We are concerned with the case where centers of component clusters and nodes are in the same geometrical space and hence keeping same dimensions.

Now consider ρ_j's belong to $X = \mathbb{R}^N$ and the distance-function d, attached to the standard norm $\|\cdot\|$ is defined on $\big([0,1] \times \mathbb{R}^N\big) \times \mathbb{R}^N$ by: $d_\Gamma((t_i, z_i), \rho_j) = \|\rho_j(t_i) - z_i\|^2$, i.e., the distance between the point z_i and mean-trajectory ρ_j.

Using some notations for a variational approach, let $A_\ell = \{\xi_i\}_{i=1}^{\ell} \subset X = \mathbb{R}^N$ be a set of graph datas to be distributed in a graph cluster $\rho = (\rho_j)_{j=1}^N$ of N component clusters ρ_j. The energy function [25] necessary to choose cluster centers of ρ_j's of a given target trajectory ρ to which elements of H_ℓ are to be assigned is the real valued function, denoted by f_ℓ, and defined on X^N by:

$$f_\ell(\rho \mid H_\ell) = \frac{1}{\ell} \sum_{i=1,\dots \ell} \bigwedge_{j=1,\dots N} d(\xi_i, \rho_j) \tag{1}$$

where $\bigwedge_{j=1,\dots M} a_j = \min\{a_1, \dots, a_N\}$ for all N real variables $a_1, \dots, a_N$ (see, e.g., [25]).

The problem here is to operate an optimal choice of the $\rho = (\rho_j)_{j=1}^N$ which minimizes the energy function $f_\ell(\cdot \mid H_\ell)$. Since this energy function $f_\ell(\cdot \mid H_\ell)$ is lower bounded and strictly decreasing, it converges absolutely to a local minimum (see [14]). We then define the general term of the real sequence of local minimum reached by $f_\ell(\cdot \mid H_\ell)$ on X^N, by $\beta_\ell = \min_{\rho \in X^N} \{f_\ell(\rho \mid H_\ell)\}$. Under certain regularity conditions and with a probability equals to 1, this sequence converges weakly for almost every sequence of observations H_ℓ.

The limit problem associated with this sequence can be defined by $\beta = \min_{\rho \in X^N} f_\infty(\rho)$. Assuming that ξ_i's are mutually i.i.d. with the same distribution $\mathbb{P}$, this allows to define the limit energy function as

$$f_\infty(\rho) = \int_X \bigwedge_{j=1,\dots N} d(x, \rho_j) \mathbb{P}(dx). \tag{2}$$

But, since $\beta_\ell \underset{\ell \to \infty}{\longrightarrow} \beta$, for almost every sequence of observations H_ℓ, we also show that with a probability equals to 1 and up to a subsequence, then minimizers ρ_ℓ of f_ℓ converge to minimizers ρ_∞ of f_∞, when $\ell \to \infty$.

We briefly outline the ideas of providing a conceptual link that are involved in obtaining such a connection between generalized mean-sets approach and extended SDA's technics. It is worth noticing that in a context devoid of turbulence phenomena generally attributed to the so called "noise", the energy function $f_\ell(\rho \mid .)$ (see [25]) and our (sampled) weight-function $\theta_{\Gamma,\ell}$, both share some similar properties under the following three crucial conditions.

4.3. Smoothing Assumptions

(*A1*) Each graph's sample signal $H_\ell = \{(t_i, h_i^{0,\ell})\}$ considered in the extended SDA's problem is isometrically attached to a data signal with finite length having as values or at its nodes, the minimizers $H_{N,0,\ell} = \{h_i^{0,\ell},\ i = 0, \dots, N-1\}$, for $\ell \geq 1$, of the weight-function of the graph. And, there is no missing or incomplete data.

(*A2*) Parameters $t_i = \mu(i) = p_i$ for $i = 0, \dots, \ell - 1$ are normalized such that $\sum_{i=0,\dots \ell-1} t_i = 1$, making each t_i a mixing scalar or weight parameter that lies between 0 and 1.

(*A3*) Consider $\rho_j(t_i) = h_j^{0,\ell}$ iff $d_\Gamma((t_i, h_i^{0,\ell}), \rho_j) = \|\rho_j(t_i) - h_i^{0,\ell}\|^2 < \epsilon$, for all $\epsilon > 0$, where d_Γ stands for the distance measure on the graph. Cluster signal assigns $\xi_j(t_i) = \rho_j(t_i) = h_j^{0,\ell}$ for $i = 0, \dots, \ell - 1$ and $j = 0, \dots, N-1$, meaning that vertices $h_i^{0,\ell}$ and $h_j^{0,\ell}$ are nearest neighbors and therefore assigned to the same center cluster ρ_j.

The next section present some challenges of providing a conceptual link of our approach with the field of cluttered environments through extended data association by smoothing problems for target tracking [2] and secure adaptive filtering [27].

5. Our Approach

Traditional k-means method is an algorithm that always starts from an initial partition of data elements into k clusters in the same observed data space [25]. It then repeatedly carries out through the so called Lloyd's iteration following two steps: first, generating a new partition by assigning each data point to the closest cluster center, and second, computing the next new cluster centers. This iteration is known to reduce the sum of squared distance of each data point to its cluster center in each iteration and thus the k-means algorithm converges to a local minimum. The new cluster centers can be easily found if the data points are in an Euclidean space [5, 25].

5.1. Generalized Mean-sets Useful Tools for Graph Signal Processing

In [7], Fotso et *al* defined the concept of chained vertex-weighted metric (VWM) graph $\mathit{\Gamma}_X$ attached to the NCCCP metric space X as an abstract triple denoted by $(V\mathit{\Gamma}_X, E\mathit{\Gamma}_X, \theta_{\mathit{\Gamma}_X})$, where $\theta_{\mathit{\Gamma}_X}$ is a weight-function on the vertex set $V\mathit{\Gamma}_X$. Such a graph typically represents dependencies between the data points through its edges. Consider N data points and let δ_{E,d_X} be the attached distance measure (e.g., Euclidean) to $\mathit{\Gamma}_X$. Each data point is considered as a node, described by a feature vector. Moreover, two data points can be connected.

Like in the general signal processing, we formalize the concept of graph signal by considering a chain (in its mathematical sense) as a dataset of N points, write

$$\begin{aligned} H_{0,\ell} &= \{\xi_i\}_{i=0}^{\ell} \\ &= \{h_j^{0,\ell} \mid j = 0, \ldots, N-1\} \subset X = \mathbb{R}^N \end{aligned}$$

of minimizers of the weight-function $\theta_{\mathit{\Gamma}_X}$. That is, a finitely many number of successive elementary minimizing nodes or stages allowing to join any given two vertices in $V\mathit{\Gamma}_X$.

We restrict ourselves to signals with a finite number N of samples and to filters with finite impulse response (FIR filters). We consider as different states of a signal, the end-vertices or nodes $h_j^{0,\ell}$, $j = 0, \ldots, N-1$, of elementary unweighted links in that chain. We also consider graph data elements as the points in space-time, the clusters as functions from time to the vertex-set, and the cluster center as the curve of best fit.

We are now able to tackle our challenge using k-means methodology on graphs, which is modeled on the corresponding k-means methodology in conventional k-means methods. In other words, a graph signal $s(t)$ is given by the vector

$$s(t) = h^{0,\ell}(t) = \left(h_0^{0,\ell}(t), \cdots, h_{N-1}^{0,\ell}(t)\right). \tag{3}$$

But, using a formal variable z^{-1}, called the shift or delay, this signal of N-samples can take the formal polynomial representation form as

$$s(z) = \sum_{j=0}^{N-1} s_j(t) z^{-j}. \tag{4}$$

Clearly, given $s(z)$, we can recover the signal s through a Fourier transform relationship. Indeed, the signal s is recovered from its Fourier coefficients by the so called inverse discrete Fourier transform [13, 19].

Let's reinterprete the finite graph signals as vectors rather than sequences or tuples as usual, and represente a graph filter as a matrix $\mathbf{A}$ called vertex-adjacency matrix of the graph or graph shift matrix. This matrix traduces the connectivities among nodes and egdes in the graph and can therefore be adopted as the shift for this entire graph. When it is defined for a given set of vertices, the so called vertex-adjacency matrix completely determined the graph, as it describes the vertices weight and connectivity. Contrary to the only edge-weighted graphs commonly involve in graph signal processing, we are dealing with graphs that are weighted on their vertices.

Let $\mathit{\Gamma}_X = (V\mathit{\Gamma}_X, E\mathit{\Gamma}_X, \theta_{\mathit{\Gamma}_X})$ be an irregular and undirected chained VWM graph attached to a dataset X (see [7]). Remember that in graph signal processing, ordering the samples corresponds to labeling the nodes of the graph, and this labelling or numbering fixes the adjacency matrix of the graph, and hence the graph shift. From a system of N given vertices of the graph $\mathit{\Gamma}_X$, the vertex-adjacency matrix of this doubly-weighted graph is a matrix, denoted by $\mathbf{A} = (a_{ij})$, of order $N \times N$ defined as:

$$a_{ij} = \begin{cases} \dfrac{\theta_{\mathit{\Gamma}_X}(x_j)}{\theta_{\mathit{\Gamma}_X}(V\mathit{\Gamma}_X)}, & \text{if } i = j \text{ and if the vertex } x_j \text{ is weighted,} \\ 1, & \text{if } i \neq j \text{ and if vertices } x_i, x_j \text{ are connected,} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

where $\theta_{\mathit{\Gamma}_X}$ stands for the weight-function defined in [7] and $\theta_{\mathit{\Gamma}_X}(V\mathit{\Gamma}_X) = \sum_{x \in V\mathit{\Gamma}_X} \theta_{\mathit{\Gamma}_X}(x)$ represents the total weight of the finite graph.

Such that the output graph signal, denoted by $s_{out}(t)$, is defined by

$$s_{out}(t) = \mathbf{A}\, s_{in}(t) \tag{6}$$

where $s_{in}(t)$ stands for the input graph signal.

Remark 5.1. Let denoted by $\mathbf{D} = diag\,(a_{11}, \cdots, a_{NN})$, the diagonal matrix where each diagonal element a_{jj} is equal to the relative weight of the vertex x_j. If $\mathbf{W}$ stands for the classical weighting matrix (with entries in $\{0; 1\}$) for the edge-weighted graph, then $\mathbf{A} = \mathbf{W} + \mathbf{D}$.

Let's assign to the linear convexification operator $\mathcal{K} : X \to X$ of the underlying metric space X its matrix of order $N \times N$, denoted by $\mathbf{K}$, called convexification matrix, relatively to a canonical base of N vectors. We simply write

$$\mathcal{K}(x) = \mathcal{K}x = \mathbf{K}x \tag{7}$$

for all $x \in X$, when there is no confusion.

In the setting of classical signal processing, convex combination approaches are already known to help improve tracking performance of adaptive filters. Indeed, Arenas-García *et al.* [2, 3] analyzed the tracking behavior of one such approach, showing that a convex combination of two adaptive LMS filters performs as close as desired to the best of its components, and possibly, better than any of them. They did it without carring about loss of input data's changes in terms of quality or quantity, i.e., the so called "informative deterministic behaviour" or "informative structure" (see, e.g., [28]).

In our context, the setting (in terms of problem-situation) and the process for filtering and clustering are different from traditional cases of signal processing, and concerns data indexed on nodes of graphs. The problem-situation here concerns target graph clustering and filtering problems under clutter with privacy.

Now, let's present more clearly the framework of our approach.

5.2. Framework of Our Approach

We consider a context where the smoothing assumptions (*A1*), (*A2*) and (*A3*) stated in Section 5 are defined similarly and are satisfied. Let's denote the time parameter by t, while noting that $t = n$ for discrete time or discrete spatial index, depending on the application.

As stated in [6], the conventional definition of an adaptive filter is based on four aspects that follows: (i) signals are processed by filters, (ii) the structure that defines how the output signal of the filter is computed from its input signal, (iii) parameters within this structure can be iteratively changed to alter the filter?s input-output relationship, and (iv) the adaptive algorithm that describes how the parameters are adjusted from one time instant to the next. In the context of graph signals, a direct extention of this concept to graphs is not straightforward, because of some differences due to the irregular nature of graphs [19].

Very often, besides signals, we also have filters. Being focused on FIR filters, remember that a FIR filter can also be represented by a polynomial in time delay or time shift denoted by z^{-1}, i.e.,

$$h(z) = \sum_{k=0}^{N-1} w_k z^{-k} \tag{8}$$

Below is an illustrative structure of an adaptive convex combination of two transversal adaptive FIR filters, in the setting of traditional signal processing.

In our case, two FIR filters with complementary capabilities adaptively mix their output signals to get an overall filter of better performance. The mixing process is carried on via a convex combination operation with mixing scalar in $\lambda(t)$.

Note that the regressor or signal vector $s(t)$ of lenght N is obtained from a random process as:

$$s^T(t) = (s(t), s(t-1), \cdots, s(t-N+1)), \tag{9}$$

where the super script $(.)^T$ denotes the transpose of a real valued vector or matrix.

The theoretical grounds of conventional signal processing in the Encrypted domain come from the field of secure function evaluations [27] and more general support generic function evaluations [22]. Remember that the existence of efficient solutions for a secure execution of some generic functions is still problematic. However, many efficient and secure technics have been developed for specific applications in the past years, building up a set of tools that foretell the potential of this technology [27].

5.2.1. Privacy-preserving Scenario Using Our Approach

This concerns the first party of our approach which can be alligned in the category of Data masking techniques. Nonetheless, there are some other examples of privacy-preserving techniques commonly used in the context of Cryptography and Data processing. Without details, we have Homomorphic encryption, Secure multi-party computation, Differential privacy, Federated learning, Data masking, Zero-knowledge proofs, etc. (see, e.g., [27]). In Graph signal theory, convexification operation refers to techniques that aim to transform a given graph signal into a convex form that is easier to analyse and manipulate. This is particularly useful in optimization problems by ensuring the convexity of the objective function, signal denoising by enforcing convex properties of signals or in graph-based learning by enhancing algorithms for tasks like clustering and filtering.

Let $H_{0,\ell} = \{\xi_i\}_{i=1}^{\ell}\}$ be a dataset to be clustered and $\{s_j = h_j^{0,\ell} \mid j = 0, \ldots, N-1\} \subset X = \mathbb{R}^N$ a sample signals indexed by nodes in the graph. Assume that the tracking is done in a two-party setting. First party, a judicious partition is operate by graph clustering and filtering through N clusters. Second party, follows a special fusion by graph fitting the first party outputs to a singletarget, i.e., an overall structure of a convex combination of two adaptive FIR filters. We understand by special fusion, a sort of fitting without a significant loss of data's qualities or properties. That is, in such a way that the privacy of input element's properties are preserved enough, from the input partition up to the output fusion. At this level, one might think at a conventional (vectorial) compression problem as the situation only involves fitting a single curve to a set of points (see 4.2).

We now describe the role of the convexification operator $\mathcal{K}$ in our method.

The convexification operator $\mathcal{K}$ as a data maskage operator. For a graph signal $s(t)$ attached to $H_{0,\ell}$, such that $s_j(t) = h_j^{0,\ell}(t)$ are the signal values at the nodes indexed respectively $j = 0, \ldots, N-1$, let's define and denote the regressor vector signal characterazing individual secured graph signal by,

$$\bar{s}(t) = \mathbf{K}\, s(t) = (\bar{s}_0(t), \cdots, \bar{s}_{N-1}(t)) \tag{10}$$

where $\bar{s}_j(t) = \mathcal{K}h_{j,\ell}(t)$ for $j = 0, \cdots N-1$, denotes the secured or masked individual input signals and $\mathcal{K}$ is the convexification operator of the underlying dataset space X, with matrix $\mathbf{K}$, standing for the secure function that ensures the privacy preserving action on individual signals. Indeed, the action of $\mathcal{K}$ on individual graph signals can seen as bringing up a sort of protection at each nodes of an input signal against loosing of some informative deterministic behaviour during the filtering process. To make simple in mathematical terms, $\bar{s}(t)$ just represents the convexifier of $s(t)$ (see, e.g., [7, 24]).

We are to study the secured weighted mean-set (expectation) scheme which obtains the output overall secured graph signal, denoted by $\bar{\mathbb{S}}(t)$, as a convex combination mean [7] of individual secured signals

$$\bar{\mathbb{S}}(t) = \mathbb{E}\bar{s}(t) = [p_j(t),\, \bar{s}_j(t)]_{j=0}^{N-1} \tag{11}$$

such that the output of the j-th component secured adaptive graph filter at time t is defined as

$$\bar{Y}_j(t) = \bar{W}_j^T(t)\bar{\mathbb{S}}(t) \tag{12}$$

for $j = 0, \cdots N-1$, where $\bar{W}_j^T(t)$ is the transposed weight vector characterazing individual secured adaptive graph filters.

Now, for general transversal secured schemes, we shall assume that

$$\bar{W}_i(t+1) = f\left(\bar{W}_i(t), \bar{s}(t), d(t), p_i(t)\right),\ i = 0, \cdots N-1 \tag{13}$$

where f represents a well-defined linear function referring to the activation or adaptation function of the output data.

The *a priori* error signal component or the so called usual noise of the individual graph filters is defined, compared to the desired signal, as $e_i(t) = \left(d(t) - \bar{W}_i(t)\right)^T \bar{s}(t)$. While the graph filter vectors $(\bar{W}_1(t), \cdots, \bar{W}_N(t))$ have length N, so that the overall secured filter can also be thought of as a secured filter with the weight vector defined as:

$$\bar{W}(t) = \left[\lambda_j(t),\, \bar{W}_j(t)\right]_{j=1}^{N}. \tag{14}$$

And the *a posteriori* error signal vector or the so called usual noise of the overall secured filter is defined, compared to the desired signal, as $e(t) = \left(d(t) - \bar{W}(t+1)\right)^T s(t)$.

Remember that, whatever the method used in convexification operation, the process typically involves altering the original signal, which can lead to loss of certain informations about its original structure. Indeed, when convexifying, one often lose the non-convex characteristics that were present before. For instance, if the signal has multiple local minima, its convexified version may only have one global minimum, making it easier to optimize.

We are dealing with privacy-preserving of signals coming from one party during their adaptive filtering by another untrusted party. In such a way that, the result of each iteration is either disclosed or given in encrypted form before the next iteration. Indeed, we are focused on the case in which component signals will be independently convexified and filtered, using their own design rules and errors, the regressor vector $s^T(t) = (1, \ldots, 1)$ of lenght N, while the mixing parameters, $p_i(t)$, will be constant and chosen to minimize the quadratic error of the overall mean-set output signal.

Filter's parameters w_j are chosen to bring the error signal $e_i(t)$ or $e(t)$ to zero or make them become minimum, i.e., such that the magnitude of the noise decreases over time.

Theoretical scheme of our processing. In our processing, we assume that the partition is conducted with no noise. Note that graph signals can be denoised. After partition by clustering has been efficiently performed and the transversal filters tracks match to targets, we then need to preserve as much as possible the "informative deterministic behaviour" or "informative structure" of these transversal filters (see, e.g., [28]). This is done by including their quantitative and qualitative properties before compressing or convexly combining them (datas) for fitting to an input overall graph filter.

Hencefore, in order to avoid track loss in such a scenario, a common approach is to use a predefined action to secure the association. Fortunately, in our case, the interpretation of the operation called data masking, data preservation or normalization can be ensured by the convexification operator $\mathcal{K}$ of the underlying data-space X.

Indeed, a cluster $x \in X$ is normalized by the convexification operator $\mathcal{K}$ to become a so called convexified or masked cluster $\mathcal{K}x \in X$ of the same dimension with x according to the linearity of $\mathcal{K}$. By normalizing, one can also understand the need and use of efficient smoothing methods that will allow to retain (i.e., to secure through masking) informative trends in the components signals or filters while disregarding noises and other undesired non-deterministic components. To make sure that the combination of input filters or graph's datas from different systems will be consistent, the input datas must first be transformed in a common domain, hereby represented by the underlying convexifiable domain, denoted by $\mathcal{K}X$.

Next, the interpretation of the fusion or compression action for a suitable fusioning can be facilitated by our convexification function Ψ_{μ_ξ}, to get the reference point data, i.e., the reference cluster properties of the input and output clusters. Then, this reference graph data point is used as the input data in the regression model [2]. This operation is done by using a data's convex combination approach, which consists of treating the subject as a problem of convex combination in a graph by using the generalized weighted mean-set approach. With this method, individual clusters like component signals that have been masked, i.e., convexified, in a previous step to become what we can call transversal convexified signals are convexly combined to obtain a new single-cluster (i.e., the overall transformed signal) represented by our so called "generalized weighted mean-set", which is then used to making the final devoted use, that is filtering.

In other words, by our approach, the individual or the so called transversal clusters are independently adapted using their own error or properties, thank to the convexificaton operator $\mathcal{K}$. While the convex combination is adapted by mean-sets of a stochastic gradient procedure in order to minimize the error of the overall structure [3], thank to our convexification function Ψ, defined as follows:

Definition 5.1. ([7]) The convexification function of X with N X-valued entries attached to a measure $\mu \in \text{Prob}_N$ with $\mu(i) = p_i$ for $i = 0, \dots N-1$, is defined to be the continuous function $\Psi_\mu : X^N \longrightarrow X$ as

$$\Psi_\mu(x_0, \dots, x_{N-1}) = [p_i, \mathcal{K}x_i]_{i=0}^{N-1}. \tag{15}$$

It is worth noticing that, our so called convexification function combines multiple points to create a new point within their convex hull, preserving the convex nature of the original set, i.e., in a way that retains their convexity. We could have just call this function, the convex-combination operator on X of lenght N. Meanwhile, the commonly called convexification operation transforms a non-convex structure (function or signal) into a convex one, often altering the original structure for optimization purposes. In essence, convexification focuses on function transformation, while convex combination deals with point aggregation within a convex space.

In general, the inverse operation of the so called convexification does not exist in a straightforward manner, because the process generally involves altering and loss of non-convex informations about the original structure. But, one can sometimes approximate, reconstruct features or recover aspects of the original structure by using some techniques like deconvexification or regularization adjustments within specific contexts.

When building a connection of our approach to SDA's technics, we dont really care about true or false tracks (see, e.g., [5, 15]) nor the optimizing choice operation of cluster centers. We just take into account the finiteness length of clusters and assumed that, as time runs forward, the k-means method favors the partition by smoothing that correctly matches tracks to clusters.

A rough illustration of how our approach shows that it theoretically matches in a two parties with data clustering and filtering processes. The first party is illustrate by the following incidence diagram (see Figure 1).

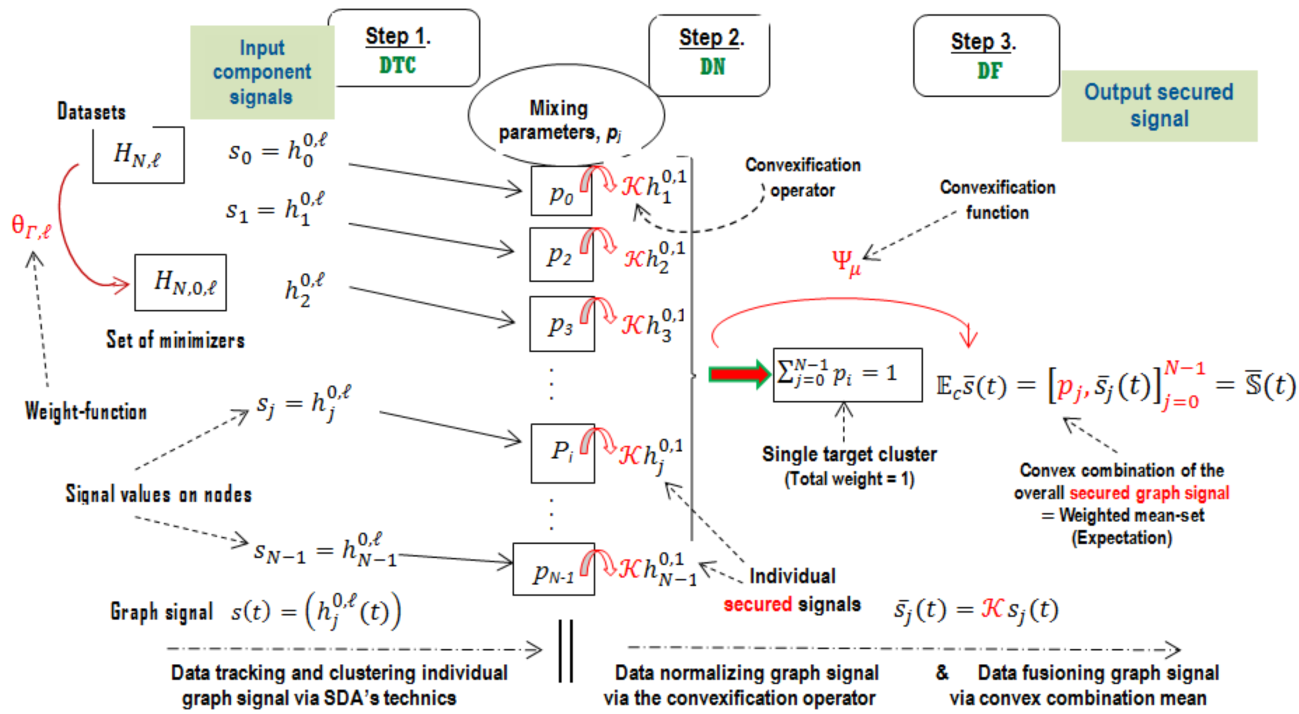

Figure 1. *Connection of our approach to graph signal processing (Part 1).*

To well understand the proposed method, we provide a short description of this Part 1 Incidence diagram, which is based on three fundamental steps.

5.2.2. Description of the Incidence Diagram (Figure 1)

In this diagram, we proceed to a sort of private data processing before filtering in the second party. Here, we have the known data (input signal) in the first step called Step 1. The middle step called Step 2 is a hidden step in which the input signal received the first transformation under the convexification operator, and the last step called Step 3 is subject of the last transformation to have the signal mean-sets (expectation) form via convex the convex combination operation. These three steps are labelled as follow:

Step 1: Data tracking and clustering (DTC) by extended SDA's technics on graphs. In this step, we start with a distribution of rough datas in an auxiliary cluster of N components using the so called moving average, which is the most popular method of smoothing a series of datas. This is a problem of smoothing data association which leads to the distribution of a finite number of data sub-groups according to the distances to cluster's centers [5], and which are easier to handle.

This first step can be consider as the step for track, partition and filtration, i.e. identifying and good partitioning of clusters, and this can be modelized by the following convergence of minimizing sets sequence:

$$H_{N,\ell} \underset{\text{minimizing}}{\longrightarrow} H_{N,0,\ell} \underset{\ell\to\infty}{\longrightarrow} H_{N,0,1}. \tag{16}$$

At this level, both the convergence and continuity properties of energy functions to carry on the smoothing partition ensure that the following equation holds

$$f_\ell(\mu_\ell \mid H_{N,0,\ell}) \underset{\ell\to\infty}{\longrightarrow} f_\infty(\mu_\infty \mid H_{N,0,1}). \tag{17}$$

Step 2: Data normalization (DN) via convexification on graphs. At this level, one wants to pass the N data subgroups through a single cluster, i.e., a singletarget, while preserving the informative deterministic behaviour carried on by each datas subgroup. That is, in such a way that in the output of the final track, we obtain datas capable of providing informations on the entire input sample (from the component filters). A condition that the generalized weighted mean-set concept for NCCCP metric spaces is up to satisfy.

At this moment, we speak of data normalization by convexification. The action we called normalizing the informative deterministic behaviour involves the use of the convexification operator $\mathcal{K}$ and after this, the convergence of set-convexifiers sequences.

Note that by Fatou's lemma, $\theta_{\Gamma,\ell}$ is a lower semicontinuous function on $\mathcal{E}_N$ that reaches its minimum on compact subsets of $\mathcal{E}_N$ (see, e.g. [14]). Under some good convergence conditions allowing to obtain strong convergence of a minimizing sequence from τ-convergence, and the continuity property of the linear operator $\mathcal{K}$, we have the convergence of adaptive secured clusters

$$\mathcal{K}H_{N,0,\ell} \underset{\ell\to\infty}{\longrightarrow} \mathcal{K}H_{N,0,1}. \tag{18}$$

Therefore, this step can then be modelized by the following formula

$$\theta_{\Gamma,\ell}(\mathcal{K}H_{N,0,\ell}) \underset{\ell\to\infty}{\longrightarrow} \theta_\Gamma(\mathcal{K}H_{N,0,1}) \tag{19}$$

since $\theta_{\Gamma,\ell}$ is continue, and where the convexifier of the set $\{a_i\}_{i=0}^{N-1}$ is defined to be the set

$$\mathcal{K}\{a_i\}_{i=0}^{N-1} = \{\mathcal{K}a_i\}_{i=0}^{N-1}. \tag{20}$$

In the next step, we move from multitarget to singletarget tracking.

Step 3: Data fitting (DF) via Convex Combination on graphs. This step is a sort of compression or fusion using particularly the convex combination technics (in the sense of Terán and Molchanov) [24], i.e. to combine and bring into a single input set, the N secured transversal input elements coming out from the data subgroup of Step 2. The acheivement of this step 3 involves the use of the convexification function defined in [7]. At this level, the mixing parameters of the smoothing and convexifying qualities is hereby expressed and traduced by probabilities $\mu(i) = p_i$ for $i = 0, \dots N-1$ of known frequencies identified by filtration/repartition with respect to some conditions or thresholds. This activity is sometimes called segmentation [28].

Since the convexification function Ψ_μ is continuous with respect to its arguments and the convexification operator $\mathcal{K}$ is linear, we have

$$\Psi_{\mu_\ell}(H_{N,0,\ell}) \underset{\ell\to\infty}{\longrightarrow} \Psi_{\mu_1}(H_{N,0,1}). \tag{21}$$

If good convergence conditions are respected as in Step 2, then Step 3 can be modelized by the following:

$$\theta_{\Gamma,\ell}(\mathcal{K}H_{N,0,\ell}) \underset{\ell\to\infty}{\longrightarrow} \theta_\Gamma(\mathcal{K}H_{N,0,1}) \tag{22}$$

which simply traduces the convergence of the empirical mean-sets to the theoretical mean-set of the entire dataset [7].

Therefore,

$$[\mu_\ell(j),\, \bar{s}_{j,\ell}]_{j=0}^{N-1} \underset{\ell\to\infty}{\longrightarrow} [\mu_1(j),\, \bar{s}_{j,1}]_{j=0}^{N-1}. \tag{23}$$

It is worth noticing that beside the convex combination process in this Step 3, each transversal component is adapted by a filter using its own properties or rules and limits or errors, while the mixing parameters, i.e. $t_j = p_j$ in our setting, are given to minimize the quadratic error of the overall data-set or structure.

5.2.3. Graph Filtering Processing Using Our Approach

This concerns the second party or our approach of processing. In the general setting of digital Signal processing, tracking properties of two adaptive finite impulse response (FIR) graph filters or secure adaptive FIR graph filter, the main objective is to study an adaptive scheme which obtains the output,

$$s_{out}(t) = y(t) = \mathbf{H}\, s_{in}(t) \tag{24}$$

with $\mathbf{H}$ standing for the overall graph filter matrix. For a convex combination of two adaptive FIR filters, the output signal of the overall structure is defined as:

$$y(t) = [\lambda(t), y_1(t); (1-\lambda(t)), y_2(t)] \tag{25}$$

where $y_j(t),\ j = 1, 2$ are the output signals of the component adaptive FIR filters defined at time t. Note that, when $\lambda(t)$ is conveniently updated in this convex combination, the resulting adaptive filter performs like the best of its individual component, or even better than any of them, under certain conditions.

Remark 5.2. In this paper, we would like to focus on shift-invariant graph filters, for which applying the graph filter $\mathbf{H}$ to the output is equivalent to applying the graph shift $\mathbf{A}$ to the input prior to filtering. Such that, the matrices $\mathbf{A}$ and $\mathbf{H}$ commute, i.e., $\mathbf{HA} = \mathbf{AH}$. But, this equality condition holds if and only if $\mathbf{H}$ is a polynomial in the graph shift $\mathbf{A}$, i.e., there exists a polynomial $h(x) = \sum_{i=0,\ldots N-1} h_k x^k$ with real or complex coefficients, such that $\mathbf{H} = h(\mathbf{A})$, and called polynomial filtering matrix.

For now, the structure of the adaptive graph filters is not considered, except for the fact that it contains adjustable parameters whose values can affect how the output signal $y(t)$ is computed. The output signal is then compared to the desired response signal $d(t)$, by subtracting the two samples signals at time t. Remember that, the main idea behind is that, if $w_i(t)$'s are assigned appropriate values at each iteration then the above scheme would extract the best properties of the component filters $y_i(t)$.

The second party of our processing, called Part 2, delivers the reference signal through the combined system model driving the filtering of the earlier secured input signal into its convex combination form. Here, the performance of a convex combination of two adaptive FIR filters on graphs is involved, having as the input signal, $\bar{\mathbb{S}}(t)$ as defined according to Equation (11), and $\bar{\mathbb{Y}}(t)$ standing for the output overall signal.

The diagram below shows the adaptive configuration.

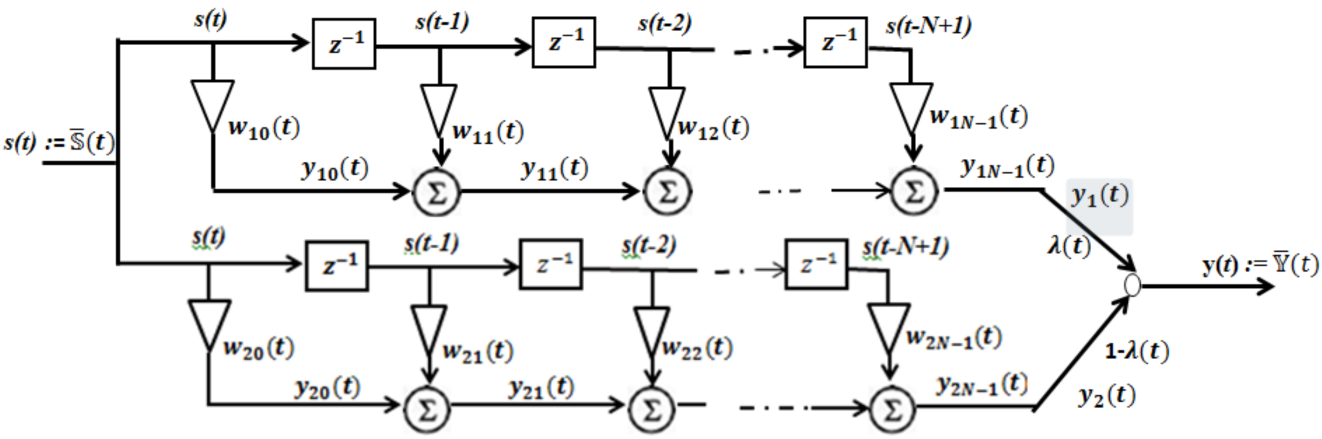

Figure 2. *Connection of our approach to graph signal processing (Part 2).*

It is worth noticing that, one nice property of a FIR graph filter is that, all the nodes share parameters among them, such that, for two nodes indexed by i and j, respective output signals are defined as convex combinations as follows:

$$y_i(t) = [w_{ik}(t), [\mathbf{A}^k s(t)]_i]_{k=0}^{N-1} \tag{26}$$

and

$$y_j(t) = [w_{jk}(t), [\mathbf{A}^k s(t)]_j]_{k=0}^{N-1} \tag{27}$$

where the shifted signal at node i is computed as a local linear combination of the signal values at neighboring nodes, i.e.,

$$[\mathbf{A}^k s(t)]_i = \sum_{j \in N_i \cup \{i\}} a_{ij} s_j \tag{28}$$

where $\mathbf{A}$ stands for the graph shift. This shows that the convolutional filter which is a shift-and-sum operation of the input signal, is defined such that the k-shifted signal $\mathbf{A}^k s(t)$ is weighted by some parameters of the same column of rank k, say $w_{.k}(t)$.

The output signal at the level of the k-shifted signal is defined by

$$y_{ik}(t) = w_{jk}^T(t) s_k(t) \tag{29}$$

for $k = 0, \cdots, N-1$ and with $w_{jk}^T(t)$ standing for the transposed weight vector characterazing individual filters. Remember that the scalar $w_{jk}(t)$ is for adapting quantitative factors also called mixing or weight parameters, lying in the real interval $[0, 1]$.

A summary of our processing is as follows: in the first party, after the DTC step where the rough graph datas have been judiciously (i.e., uniformly in our context) partinioned in N individual cluster signals, follows the step of normalizing datas (DN) by convexification. At this level, all N transversal output datas of the DTC step have to be secured and judiciously inserted into a next singletarget, after being transformed using $\mathcal{K}$ operator. The process continues with the third step called data fusion (DF) using Terán and Molchanov's convex combination operation. At this level, the convex combination mean, i.e., the generalized mean-set (expectation) of these N secured transversal output coming from DN step is computed and represents the new cluster input in part 2, to be fitting into an overall convex combination of two complementary capabilities adaptive FIR filters.

6. Result: Our Model

The private and last output scenario is more realistic, and it is the one on which we are fucused in this paper, as it corresponds to the case where the FIR block, i.e., our weighted mean-set (expectation) adaptive filter, can be used as a module of a more complex system whose intermediate signals must not be disclosed to any party (see, [27] for the case of LMS filters). Our theoretical and experimental diagram highlight the utility of our framework over previous approaches in settings in which the measurements available are unnoisy or small in number.

To move from a signal in the graph node domain to the spectral frequency domain and come back, we use a tool like the graph Fourier transform.

6.1. Graph Fourier Transform

From the eigendecomposition of $\mathbf{A}$ [19, 29], the graph Fourier transform $\mathbf{F}$ is the inverse of the matrix $\mathbf{V}$ of eigenvectors of the shift $\mathbf{A}$,

$$\mathbf{F} = \mathbf{V}^{-1} \tag{30}$$

But, the graph Fourier transform of graph signal s is given by the graph Fourier analysis decomposition

$$\hat{s}(t) = \mathbf{F}\, s(t) = \{\hat{s}_k(t) \mid k = 0, \ldots, N-1\} \tag{31}$$

where

$$\hat{s}_k(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} s_n(t) e^{-j\frac{2\pi}{N}kn} \tag{32}$$

stand for the Fourier coefficients of the signal.

Coming back to Equation (25) concerning ordinary convex combination of two adaptive FIR filters, when generalizing it to a sample of N individual secured adaptive graph filters, the final secured output of Part 2.

6.2. Modelization of the Secure Adaptive Graph Filtering

The overall weighted-mean secured output graph signal, denoted by $\bar{\mathbb{Y}}(t)$, is obtained from the weighted-mean graph signal output of Part 1, denoted by $\bar{\mathbb{S}}(t)$, using the Equation in (12), shift-invariance of matrices and linearity of outputs, as

$$\begin{aligned} \bar{\mathbb{Y}}(t) &= [\lambda(t), \bar{Y}_1(t); (1-\lambda(t)), \bar{Y}_2(t)] \\ &= [\lambda(t), \bar{W}_1^T(t)\hat{\bar{\mathbb{S}}}(t); (1-\lambda(t)), \bar{W}_2^T(t)\hat{\bar{\mathbb{S}}}(t)] \\ &= [\lambda(t), [\bar{w}_{1k}(t), [\mathbf{A}^k\hat{\bar{s}}(t)]_1]_{k=0}^{N-1}; (1-\lambda(t)), \\ &\quad [\bar{w}_{2k}(t), [\mathbf{A}^k\hat{\bar{s}}(t)]_2]]_{k=0}^{N-1} \\ &= [\lambda(t)\bar{w}_{1k}(t), [\mathbf{A}^k\hat{\bar{s}}(t)]_1; (1-\lambda(t))\bar{w}_{2k}(t), \\ &\quad [\mathbf{A}^k\hat{\bar{s}}(t)]_2]]_{k=0}^{N-1} \\ &= [\lambda(t), \mathbf{H}_1\mathbf{A}; (1-\lambda(t)), \mathbf{H}_2\mathbf{A}]\, \hat{\bar{s}}(t) \\ &= [\lambda(t), \mathbf{H}_1; (1-\lambda(t)), \mathbf{H}_2]\, \mathbf{AFK}s(t). \end{aligned} \tag{33}$$

Therefore, the graph vertex-weighted adjacency matrix of the overall adaptive convex combination of two FIR filters takes the form:

$$\mathbf{H} = [\alpha_1, \mathbf{H}_1; \alpha_2, \mathbf{H}_2]\, \mathbf{AFK} \tag{34}$$

where $\alpha_i,\ i = 1,2$ are non-negative coefficients such that $\alpha_1 + \alpha_2 = 1$, stand for the mixing parameters, matrices $\mathbf{H}_i,\ i = 1,2$ stand for individual FIR filters, matrix $\mathbf{A}$ stands for the graph shift, $\mathbf{F}$ stands for the Fourier transform and $\mathbf{K}$ stands for the convex combination operator of the underlying dataset.

Remark 6.1. At the end of the process, the output signal can be recovered using the inverse graph Fourier transform which computes the output back in the graph node domain. [11, 13, 19].

6.3. Description of Our Model

According to (33), secure adaptive graph filtering process by $\mathbf{H}$ can be performed by firstly convexifying the input signal components $\mathbf{K}s(t)$ in the graph node domain. After what, the graph Fourier transform of these convexifiers output is obtained via $\mathbf{FK}s(t)$; next followed by shift multiplication in the frequency domain of the graph by the filter frequency response $\mathbf{AFK}s(t)$. Finally, the graph filtering operation is done by the convex combination of the two component graph filters.

Our model in Equation (34), is a composed transformation made up of a convolutional graph filter $[\alpha_1, \mathbf{H}_1; \alpha_2, \mathbf{H}_2]\, \mathbf{A}$ and the matrix product $\mathbf{FK}$. This represents the fundamental matrix model describing the entire idea developped in this paper. The output overall secured graph signal can now be filtered in its better form to a preferable graph cluster.

But, this study involves adjusting graph filters based on incoming graph data. Therefore, our resulting model looks more flexible as it addresses security in data transmission or processing, focussing on protecting data or information from unauthorized access or manipultaion. It also adapts changes in the input signal, which is crucial for effective performance. Compares to some others approaches, this model enables dynamic adaptation based on statistcal properties of the input graph signals allowing for more flexible adaptation mechanism, like filtering adjustments.

To implement the performance of our proposed model, we examine the following theoretical case study.

7. Theoretical Case Study

We start by assuming that the two transversal FIR filters $y_1(t)$ and $y_2(t)$ considered, are of the same structure, i.e., with the same step-size, constant mixing or weight parameters liying in the interval $[0,1]$, metrised by any ideal probability measure of metric d_Γ of order $\alpha = 1$, and the optimal solution not subject to changes at different speeds. From the example above, the convexification operator $\mathcal{K}$ coincides in this case with the identity map of the convexifiable space $\mathbb{R}^2$.

We are to study the stationary performance of the adaptative weighted mean-set (expectation) scheme which obtains the output of the overall filter according to

$$\bar{\mathbb{Y}}(t) = \left[\lambda_i(t),\, \bar{Y}_i(t)\right]_{i=0}^{1},$$

where $\bar{Y}_i(t)$ are defined to be $\bar{Y}_i(t) = \bar{W}_i^T(t)\bar{\mathbb{S}}(t),\ i = 1,2$, with $s^T(t) = (1,1)$. Here, the mixing parameters $\lambda_i(t) = p_i(t) = p_i \in [0,1]$ are constant.

We refered to the example study in [29], illustrating a convex combination approach for Artificial neural network, to consider some similar numerical and fonctional parameters. For the case of linear structure, we considered the same adaptation or activation function, defined by $y = f(x) = 1+5x$, where x stands for the input variable (i.e., the reference point data used as the input data in the regression model) and y stands for the output variable (i.e., the overall best reference output response) after secure adaptive filtering.

We theoritically study two illustration scenarios of two data generation processes which are similar in structure (linear), with different weight parameters respecting condition ($A2$) (see 5) and without noise. We use linear regression based on convex combination method to determine the best reference of the convex combination adaptive secured graph filters.

In this example, two illustration scenarios are traduced under privacy as follows:

Scenario 1: We considered equal mixing parameters: $p_0 = p_1 = 0.5$

We have

$$\bar{\mathbb{Y}}(t) = 0.5\bar{Y}_0(t) + 0.5\bar{Y}_1(t) = 1 + 5(0.5\bar{s}_0(t) + 0.5\bar{s}_1(t)).$$

Scenario 2: We considered different mixing parameters: $p_0 = 0.2$ and $p_1 = 0.8$

We have

$$\bar{\mathbb{Y}}(t) = 0.2\bar{Y}_0(t) + 0.8\bar{Y}_1(t) = 1 + 5(0.2\bar{s}_0(t) + 0.8\bar{s}_1(t)).$$

For each of these scenarios, after performing about 100 replications using an Intel Core i5-6300U CPU @ 2.40GHz, reporting root mean squared errors signals, high performances were still observed for our model in scenario 2. That is, $3.469(1.322)$ for scenario 1 and $2.978(1.003)$ for scenario 2, compared to $3.4723(1.3234)$ for scenario 1 and $2.9894(1.0022)$ for scenario 2, obtained in [29].

The obvious similarities in these results can find a justification in the fact that, the convexification operator $\mathcal{K}$ of the data-space which is supposed to bring some privacy changes by the convexifying action coincides with the identity function of the data-space.

Our approach may face some two practical limitations in certain aspects like Dynamic changes or data quality, since many graph methods in the literature assume static structures, while real-world networks often change dramatically (e.g., removing or adding nodes and edges). And, also like limited focus as many approaches may assume homogeneity in node or edge characteristics, which can be unrealistic in diverse networks. These limitations highlight the challenges and considerations when applying Mosina's approach and its generalized versions in practical scenarios, emplasizing the need for careful evaluation and adaptation to specific contexts.

8. Conclusion

The generalized Mean-sets theory's approach developped from Mosina's arguments on NCCCP spaces can, under some crucial conditions, help to improve adaptive graph filtering under privacy of some individual informative trends. In this paper, we have analyzed the graph filtering behaviors of an adaptive convex combination of N individual adaptive filters using this approach. Individual adaptive graph filters being independently adapted by space-valued random variables on which the natural convexification operator of the underlying metric space acts to preserve privacy of their informative trends. While the empirical weighted mean-set obtained using the convexification function on the underlying dataset (as metric space) is adapted to minimize the error signal of the overall structure. We realized that the matrix model from the overall graph signal and filter provides a generalized weighted mean-set (expectation) of the transversal secured signal and adaptive graph filters, and these performs as close as desired to the best of their components. Thanks to the properties of some generalized Mean-sets theory's tools like the expectation operator, the convexification operator and the convex combination function.

Abbreviations

CC	Convex Combination
DF	Data Fusioning (Signal)
DN	Data Normalizing (Signal)
DTC	Data Tracking and Clustering
FIR	Finite Impulse Response (Filter)
GSP	Graph Signal Processing
i.i.d.	Independent and Identically Distributed
LMS	Least Mean Squares (Filter)
MST	Mean-sets Theory
NCCCP	Negatively Curved Convex Combination Polish (Space)
SDA	Smoothing Data Association
SPED	Signal Processing in the Encrypted Domain
VWM	Vertex-weighted Metric (Graph)

ORCID

0000-0002-0383-4751 (Christophe Fotso)
0000-0003-4779-5800 (Duplex Elvis Houpa Danga)
0009-0008-2647-7428 (Daniel Tieudjo)

Acknowledgments

The authors thank the anonymous reviewers for their constructive comments and suggestions, which significantly improved the final version of the manuscript.

Conflicts of Interest

The authors declare they have no competing interests.

References

[1] A. Ahidar-Coutrix, T. Le Gouic and Q. Paris. (2019). On the rate of convergence of empirical barycentres in metric spaces: curvature, convexity and extendible geodesics. HAL Id: hal-01810530. Available at https://hal.archives-ouvertes.fr/hal-01810530v3

[2] J. Arenas-García, A. R. Figueiras-Vidal and A. H. Sayed. (2005). Tracking properties of a convex combination of two adaptive filters, IEEE/SP 13th Workshop on Statistical Signal Processing, 109–114.

[3] J. Arenas-García, A. R. Figueiras-Vidal and A. H. Sayed. (2006). Mean-square performance of a convex combination of two adaptive filters. *IEEE transactions on signal processing*, 54(3), 1078–1090.

[4] A. Bator and W. Ziéba. (2009). On some definition of expectation of random element in metric space. *Annales Universitatis Mariae Curie-Sklodowska*, 63(1), 39.

[5] C.-S. Chang, W. Liao, Y.-S. Chen and L.-H. Liou. (2016). A mathematical theory for clustering in metric spaces. *IEEE Transactions on Network Science and Engineering*, 3(1), 2–16.

[6] Douglas S. C. (1999). Introduction to Adaptive Filters in *Digital Signal Processing Handbook* by Ed. Vijay K. Madisetti and Douglas B. Williams, CRC Press LLC.

[7] C. Fotso, E. D. Houpa Danga and D. Tieudjo. (2024). An Approach of Mean-sets Theory for Negatively Curved Convex Combination Polish (NCCCP) Metric spaces, *Advances in Mathematics: Scientific Journal*, 13(02), 145-180.

[8] M. Fréchet. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré*, 10(4), 215–310.

[9] C. E. Ginestet, A. Simmons and E. D. Kolaczyk. (2012). Weighted Fréchet means as convex combinations in metric spaces: properties and generalized median inequalities. *Statistics and Probability Letters*, 82(10), 1859–1863.

[10] W. Herer. (1992). Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature. *Proba. Math. Statist*, 13(1), 59–70.

[11] E. Isufi, F. Gama, D. I. Shuman and S. Segarra. (2024). Graph Filters for Signal Processing and Machine Learning on Graphs. *IEEE Transactions on Signal Processing*.

[12] D. Jin, J. Chen, and J. Chen. (2018). Convex combination of diffusion strategies over distributed networks. *Proc. Asia-Pacific Signal Inf. Process. Association*, 224-228.

[13] M. Junaid. (2022). Signal Processing on Graphs: a state of art. HAL Id: hal-03813429.

[14] J. Lember. (2003). On minimizing sequences for k-centres. *Journal of Approximation Theory*, 120(1), 20–35.

[15] S. Memon, T. L. Song and T. H. Kim. (2016). Smoothing data association for target trajectory estimation in cluttered environments. *Eurasip Journal on Advances in Signal Processing*, 2016(21), 1–13.

[16] N. Mosina. (2009). Probability on graphs and groups: theory and applications, Ph D. Thesis, Columbia University. Available at https://www.math.columbia.edu/thaddeus/theses/2009/mosina.pdf

[17] N. Mosina and A. Ushakov. (2010). Mean-set attack: cryptanalysis of Sibert *et al.* authentication protocol. *Journal of Mathematical Cryptology*, 4(2), 149–174.

[18] N. Mosina and A. Ushakov. (2011). Strong law of large numbers on graphs and groups. *Groups Complexity Cryptology*, 3(1), 67–103.

[19] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura and P. Vandergheynst. (2018). Graph Signal Processing: Overview, Challenges, and Applications. *Proceedings of the IEEE* — Vol. 106(5), 808-828.

[20] D. Pollard. (1981). Strong consistency of k-means clustering. *The Annals of Statistics*, 135–140.

[21] D. Pollard. (1982). Quantization and the method of k-means. *IEEE Transactions on Information theory*, 28(2), 199–205.

[22] Y. S. Soh. (2019). Fitting convex sets to data: Algorithms and Applications. Ph.D Thesis, California Institute of Technology, 193 pages.

[23] K.-T. Sturm. (2003). Probability measures on metric spaces of nonpositive curvature, Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces. Lecture Notes from a Quarter Program on Heat Kernels, Random Walks, and Analysis on Manifolds and Graphs: April 16-July 13, 2002, Emile Borel Centre of the Henri Poincaré Institute, 338-357.

[24] P. Terán and I. Molchanov. (2006). The law of large numbers in a metric space with a convex combination operation. *Journal of theoretical probability*, 19(4), 875–898.

[25] M. Thorpe, F. Theil, A. M. Johansen and N. Cade. (2015). Convergence of the k-means minimization problem using Γ-convergence. *SIAM Journal on Applied Mathematics*, 75(6), 2444–2474.

[26] N. Tran Thuan. (2015). An approach for metric space with a convex combination operation and applications. arXiv e-prints, arXiv–1501.

[27] J. R. Troncoso-Pastoriza and F. Pérez-González. (2011). Secure adaptive filtering. *IEEE Transactions on Information Forensics and Security*, 6(2), 469-485.

[28] J. Wodecki, A. Michalak and P. Stefaniak. (2018). Review of smoothing methods for enhancement of noisy data from heavy-duty LHD mining machines. E3S Web of Conferences, 29(00011).

[29] W. Yamaka, R. Phadkantha and P. Maneejuk. (2021). A Convex Combination Approach for Artificial Neural Network of Interval Data. *Appl. Sci.* 11, 3997.