Differentially Private MIMO Filtering for Event Streams and Spatio-Temporal Monitoring

Jerome Le Ny and Meisam Mohammady

Abstract—Many large-scale systems such as intelligent transportation systems, smart grids or smart buildings collect data about the activities of their users to optimize their operations. In a typical scenario, signals originate from many sensors capturing events involving these users, and several statistics of interest need to be continuously published in real-time. Moreover, in order to encourage user participation, privacy issues need to be taken into consideration. This paper considers the problem of providing differential privacy guarantees for such multi-input multi-output systems operating continuously. We show in particular how to construct various extensions of the zero-forcing equalization mechanism, which we previously proposed for single-input single-output systems. We also describe an application to privately monitoring and forecasting occupancy in a building equipped with a dense network of motion detection sensors, which is useful for example to control its HVAC system.

I. INTRODUCTION

During the past few years we have witnessed a fast growing trend towards equipping our environment with swarms of sensors and actuators in order to improve the efficiency of traditional activities and industries, from healthcare to traffic management systems to power grids. At the same time, it is becoming increasingly clear that these systems pose a risk to our privacy [1], and that more work is needed to provide rigorous ways to balance privacy and utility. Privacy-utility tradeoffs can be studied under various definitions of privacy, e.g., \(k\)-anonymity [2] or information-theoretic privacy [3], but in recent years, the notion of differential privacy has been particularly popular for applications involving sensitive statistical databases [4], [5]. The main idea is that a differentially private mechanism should publish information about a dataset in a way that is not too sensitive to the presence or absence of a single individual’s data. As a result, the fact that an individual contributes its data or not does not significantly change the capability of an adversary to make inferences about that individual. One of the main practical reasons why differentially privacy has proved to be a particularly convenient notion of privacy is that it does not require modeling the side information available to adversaries to provide guarantees, even though this side information is the main reason behind most privacy breaches [6].

In this paper we discuss the application of the notion of differential privacy in the context of real-time systems, following our initial work in, e.g., [7]–[9]. In particular, these papers consider a scenario inspired by [10], [11] where a system receives a sensitive input signal recording events related to individual users, and continuously publishes a statistic of interest derived from this signal, e.g., a moving average of the number of recent events. Here we extend our analysis to multi-input multi-output systems, which considerably broadens the applicability of these ideas to common situations where multiple sensors monitor an environment and we wish to concurrently publish several statistics of interest. A typical application example is that of analyzing spatio-temporal records provided by networks of simple counting sensors, e.g., motion detectors in buildings or loop inductors in traffic information systems [12].

The rest of the paper is organized as follows. Section II introduces our problem formulation, as well the necessary background on differential privacy. Section III presents certain preliminary calculations necessary to develop our differentially private dynamic mechanisms, which are described in Section IV. Finally, in Section V we briefly discuss an application to privately estimating and forecasting occupancy in a building equipped with a network of motion detectors. Throughout the paper we use the following abbreviations: SISO for Single-Input Single-Output, SIMO for Single-Input Multiple-Output, MIMO for Multiple-Input Multiple-Output, and \([m] := \{1, \ldots, m\}\).

II. PROBLEM STATEMENT

A. Generic Scenario

We consider \(m\) sensors detecting events, with sensor \(i\) producing a real-valued discrete-time scalar signal \(\{u_{i,t}\}_{t \geq 0}\), for \(i \in [m]\). In a building monitoring scenario for example, the sensors could be motion detectors distributed at various locations and polled at regular intervals, with \(u_{i,t} \in \mathbb{Z}\) the number of detected events reported for period \(t\). We denote \(u\) the resulting vector valued signal, i.e., \(u_t \in \mathbb{R}^m\). A linear time-invariant (LTI) filter \(F\), with \(m\) inputs and \(p\) outputs, takes input signals \(u\) from the sensors and publishes output signals \(y\) of interest, with \(y_t \in \mathbb{R}^p\). For example, we might be interested in continuously updating a real-time estimate of the number of people in various parts of the building, as well as short- and medium-term occupancy forecasts, in order to optimize the operations of the Heating, Ventilation, and Air Conditioning (HVAC) system. The problem considered in this paper consists in replacing the filter \(F\) by a system processing the input \(u\) and producing a signal \(\tilde{y}\) as close as possible to the desired output \(y = Fu\) (here, in the mean-squared sense), while providing some privacy guarantees to the users from which the input signals \(u\) originate.
B. Differential Privacy

The formal definition of privacy that we adopt here is that of differential privacy, see Definition 1 below. In the previous building monitoring example, one goal of a privacy constraint could be to provide guarantees that an individual user cannot be tracked too precisely from the published data, see [13].

1) Adjacency Relation: Formally, we start by defining an symmetric binary relation, denoted Adj, on the space of datasets $D$ of interest, which is used to define what it means for two datasets to differ by the data of a single individual. Here, $D := \{ u : \mathbb{N} \to \mathbb{R}^m \}$, and we have $\text{Adj}(u, u')$ if and only if we can obtain the signal $u'$ from $u$ simply by adding or subtracting the events corresponding to one user. Motivated again by our application to motion detection, we consider in this paper the following adjacency relation

$$\text{Adj}(u, u') \iff \forall i \in [m], \exists t_i \in \mathbb{N}, \alpha_i \in \mathbb{R}, \text{s.t. } u'_i - u_i = \alpha_i \delta_{t_i}, |\alpha_i| \leq k_i. \quad (1)$$

In other words, a single individual can affect each input signal component at a single time (here $\delta_{t_i}$ denotes the discrete impulse signal with impulse at $t_i$), and by at most $k_i \in \mathbb{R}$. Note that this adjacency relation extends the one considered in [8], [10], [11] to the multi-dimensional case. The following notation will be used in the following. We denote by $k \in \mathbb{R}^m$ the vector with components $k_i$. Also, let $e_i \in \mathbb{R}^m$ be the $i$th basis vector, i.e., $e_{ij} = \delta_{ij}, j = 1, \ldots, m$. Then for two adjacent signals $u, u'$, we have

$$u' - u = \sum_{i=1}^m \alpha_i \delta_{t_i} e_i, \quad (2)$$

The adjacency definition (1) puts two constraints on the influence that an individual can have on the input data in order for our differentially private mechanisms to offer him guarantees. First, any given sensor can report an event due to the presence of the individual only once. This is a sensible constraint in applications like traffic monitoring with fixed motion detectors activated only once by each car traveling along a road, or for certain location-based services where a customer would check-in say at most once per day at each visited store. For a building monitoring scenario however, a single user could trigger the same motion detector several times over the time interval of interest for our analysis. One solution consists in splitting the data stream of problematic sensors into several successive intervals, each considered as the signal from a new virtual sensor, so that an individual’s data is present only once in each interval. However, increasing the number of inputs degrades the privacy guarantees or the output quality that we can provide. Hence in general no privacy guarantee will be offered to users who activate the same sensor too frequently. Second, we bound the magnitude of an individual’s contribution by $k_i$, but this is not really problematic in applications such as motion detection, where we can typically take $k_i = 1$. Moreover, we can place additional constraints on $k$ to capture additional knowledge about the problem, which can help design mechanisms with better performance. For example, if a given person can activate at most $l < m$ sensor and each $k_i$ is at most 1, we can add the constraint $\|k\|_1 \leq l$.

2) Definition of Differential Privacy: First, a mechanism $M$ is simply a function that takes datasets in $D$ as inputs and such that $M(d)$ is a random output signal.

Definition 1: Let $D$ be a space equipped with a symmetric binary relation denoted $\text{Adj}$, and let $(R, M)$ be a measurable space. Let $\epsilon, \delta \geq 0$. A mechanism $M$ with (random) outputs in $R$ is $(\epsilon, \delta)$-differentially private for $\text{Adj}$ if for all $d, d' \in D$ such that $\text{Adj}(d, d')$, $P(M(d) \in S) \leq e^\epsilon P(M(d') \in S) + \delta, \forall S \in \mathcal{M}$. \quad (3)

This definition quantifies the allowed deviation for the output distribution of a differentially private mechanism, when a single individual is added or removed from a dataset. In this paper, the space $D$ was defined as the space of input signals, and the adjacency relation considered is (1). The output space $R$ is simply the space of output signals $R := \{ y : \mathbb{N} \to \mathbb{R}^p \}$. Finally, a differentially private mechanism will consist of a system approximating our MIMO filter of interest $F$, as well as a source of noise necessary to randomize the outputs and satisfy (3) [8].

3) Sensitivity: Enforcing differential privacy can be done by randomly perturbing the published output of a system, at the price of reducing its utility or quality. Hence, we should evaluate as precisely as possible the amount of noise necessary to make a mechanism differentially private. For this purpose, the following quantity plays an important role.

Definition 2: The $\ell_2$-sensitivity of a system $G$ with $m$ inputs and $p$ outputs with respect to the adjacency relation $\text{Adj}$ is defined by

$$\Delta^m_p G = \sup_{\text{Adj}(u, u')} \|Gu - Gu'\|_2 = \sup_{\text{Adj}(u, u')} \|G(u - u')\|_2, \quad (4)$$

where by definition $\|Gv\|_2^2 = \sum_{k=-\infty}^{\infty} |Gv_k|^2$ and $|x| = \left( \sum_{k=-1}^{p} |x_k|^2 \right)^{1/2}$ is also used throughout the paper to denote the Euclidean norm for $x$ in $\mathbb{R}^p$ or $\mathbb{C}^p$.

4) A Basic Differentially Private Mechanism: The basic mechanism of Theorem 1 below (see [8]), extending [14], can be used to answer queries in a differentially private way. To present the result, we recall first the definition of the $Q$-function $Q(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} du$. Now for $\epsilon > 0, \delta \geq 0$, let $K = Q^{-1}(\delta)$ and define $\kappa_{\delta, \epsilon} = \frac{1}{\epsilon} (K + \sqrt{K^2 + 2\epsilon})$.

Theorem 1: Let $G$ be a system with $m$ inputs and $p$ outputs, and with $\ell_2$-sensitivity $\Delta^m_p G$ with respect to an adjacency relation $\text{Adj}$. Then the mechanism $M(u) = Gu + w$, where $w$ is a $p$-dimensional Gaussian white noise with covariance matrix $\kappa_{\delta, \epsilon}^2 (\Delta^m_p G)^2 I_p$, is $(\epsilon, \delta)$-differentially private for $\text{Adj}$.

The mechanism $M$ described in Theorem 1, producing a differentially-private version of a system $G$, is called an output-perturbation mechanism. The amount of noise introduced by this mechanism is proportional to the $\ell_2$-sensitivity of the filter and to $\kappa_{\delta, \epsilon}$, which can be shown to behave as $O(\ln(1/\delta))^{1/2}/\epsilon$. We need to add noise proportional to the sensitivity of the whole filter $G$ on each output, even if $G$ is diagonal say, otherwise trivial attacks that simply average
a sufficient number of outputs could potentially detect the presence of an individual with high probability.

In conclusion we could obtain a differentially private mechanism for our original problem by simply adding a sufficient amount of noise to the output of our desired filter \( F \), provided we can compute its sensitivity, which is the topic of the next section. However, it is possible in general to design mechanisms with much less overall noise than this output-perturbation scheme, as discussed in Section IV.

### III. Sensitivity Calculations

For the following sensitivity calculations, the \( H_2 \) norm of an LTI system plays an important role. Recall that for a system with \( M \) inputs \( \| G \|_2 = \sum_{i=1}^{m} \| G \delta_0 e_i \|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \| G(e^{j\omega}) \| G(e^{j\omega}) \| d\omega \). Note from the frequency domain definition that writing \( G(z) = \{ G_{ij}(z) \}_{i,j} \) for the \( p \times m \) transfer matrix, we have \( \| G \|_2^2 = \sum_{i,j} \| G_{ij} \|_2^2 \).

**A. Exact solutions for the SIMO and Diagonal Cases**

The following theorem immediately generalizes the SISO scenario considered in [7], [8] to the SIMO case.

**Theorem 2 (SIMO LTI system):** Let \( G \) be a stable LTI system with one input and \( p \) outputs. For the adjacency relation (1), we have \( \Delta^2 \nabla P = k_1 \| G \|_2^2 \), where \( \| G \|_2 \) is the \( H_2 \) norm of \( G \).

For a system \( G \) with multiple inputs, the special case where \( G \) is diagonal, i.e., its transfer matrix is \( G(z) = \text{diag}(G_{11}(z), \ldots, G_{mm}(z)) \) also leads to a simple sensitivity result. Note that in this case, we have \( \| G \|_2^2 = \sum_{i=1}^{m} \| G_{ii} \|_2^2 \).

**Theorem 3 (Diagonal LTI system):** Let \( G \) be a stable diagonal LTI system with \( m \) inputs and outputs. For the adjacency relation (1), we have \( \Delta^2 \nabla G = \| G \|_2^2 = (\sum_{i=1}^{m} \| k_i G_{ii} \|_2^2)^{1/2} \), where \( K = \text{diag}(k_1, \ldots, k_n) \).

**Proof:** If \( G \) is diagonal, then for \( u \) and \( u' \) adjacent, we see from (2) that \( \| G(u - u') \|_2^2 = \sum_{i=1}^{m} |\alpha_i G_{ii} \delta_i e_i |^2 = |\alpha_i k_i G_{ii} \delta_i e_i |^2 = \| G_{ii} \|_2^2 \| \delta_i \|_2^2 \). Hence \( \| G(u - u') \|_2^2 = \sum_{i=1}^{m} \| k_i G_{ii} \|_2^2 \| \delta_i \|_2^2 = \| G \|_2^2 \). The bound is attained if \( |\alpha_i| = k_i \) for all \( i \).

**B. Upper and Lower Bound for the general MIMO Case**

For MISO or general MIMO systems, the sensitivity calculations are no longer so straightforward, because the impulses on the various input channels, obtained from the difference of two adjacent signals \( u, u' \), all possibly influence any given output. Still, the following result provides a simple bound on the sensitivity.

**Theorem 4:** Let \( G \) be an LTI system with \( m \times p \) transfer matrix \( G(z) = [G_1(z), \ldots, G_m(z)] \) (i.e., with columns \( G_k \)), such that \( \| G \|_2 < \infty \). For the adjacency relation (1), we have \( \| G \|_2 \leq \Delta^2 \nabla G \leq |k| \| G \|_2 \), where \( K = \text{diag}(k_1, \ldots, k_m) \) and \( |k| = \left( \sum_{i=1}^{m} k_i^2 \right)^{1/2} \).

**Proof:** We have \( \| G(u - u') \|_2 = \sum_{i=1}^{m} |\alpha_i G_{ii} \delta_i e_i |^2 \), and thus \( \| G \|_2 = \sum_{i=1}^{m} \| G_{ii} \|_2^2 \). Hence \( \| G(u - u') \|_2 \leq \sum_{i=1}^{m} |\alpha_i k_i G_{ii} \delta_i e_i |^2 \leq \| G \|_2^2 \| \delta_i \|_2^2 \). For the upper bound, we can use the Cauchy-Schwarz inequality:

\[
\| G(u - u') \|_2^2 = \sum_{i=1}^{m} |\alpha_i G_{ii} \delta_i e_i |^2 \leq \sum_{i=1}^{m} |\alpha_i| \| G_{ii} \|_2 \| \delta_i \|_2^2 \leq \| k \| \left( \sum_{i=1}^{m} \| G_{ii} \|_2 \| \delta_i \|_2^2 \right)^{1/2} \]
Theorem 5: Let $G$ be a stable LTI system with $m$ inputs and $p$ outputs, and state space representation as described above. Then, for the adjacency relation (1),

$$\left(\Delta^{m,p}_2 G\right)^2 = \|GK\|_2^2 + \sum_{i,j=1 \atop i \neq j}^m k_i k_j \left(\sup_{t_1, t_2 \in \mathbb{N}} |\xi_{t_1-t_2}| \right).$$  \hspace{1cm} (5)

In the expression (5) for the sensitivity, the maximization over inter-event times $t_1 - t_2$ still needs to be performed and depends on the parameters of the specific system $G$. This result could be used to evaluate carefully the amount of noise necessary in an output perturbation mechanism, but unfortunately it seems too unwieldy at this point to be used in more advanced mechanism optimization schemes, such as the one discussed in the next section.

IV. ZERO-FORCING MIMO MECHANISMS

Using the sensitivity calculations above, we can now design differentially private mechanisms to approximate a given filter $F$, as discussed in Section II-A. The mechanisms described below generalize to the MIMO case some ideas introduced in [7]. Indeed, the general approximation architecture considered, described on Fig. 1, is the same as for the SISO case. On this figure, the system $H$ is of the form $H = FL$, with $L$ a left inverse of the pre-filter $G$. We call the resulting mechanisms Zero-Forcing Equalization (ZFE) mechanisms. The goal is to design $G$ (and hence, $H$) so that the Mean-Squared Error (MSE) between $y$ and $\hat{y}$ on Fig. 1 is minimized. In order to obtain a differentially private signal $v$, the variance of the Gaussian white noise signal $w$ is proportional to the sensitivity of the filter $G$. It was shown in [7] that this setup can allow significant performance improvements compared to the output-perturbation mechanism. Note that the latter can be recovered when $G = F$ and $H$ is the identity.

A. SIMO system approximation

First, let us assume that $F$ is a SIMO filter, with $p$ outputs. Note that this scenario is considered in [16] (from a very different point of view) for the special case where each row of $F$ is a moving average filter with a different size for the averaging window. Consider a first stage $G(z) = \text{col}(G_1(z), \ldots, G_q(z))$ taking the input signal $u$ and producing $q$ intermediate outputs that must be perturbed. The second stage is taken to be $H = FL$, with $L(z) = [L_1(z), \ldots, L_q(z)]$ a left-inverse of $G$, i.e., satisfying

$$\sum_{i=1}^q L_i(z)G_i(z) = 1.$$  \hspace{1cm} (6)

Let us also define the transfer functions $M_i$, $i = 1, \ldots, q$, such that $M_i(z) = L_i(z^{-1})$, hence $M_i(e^{j\omega}) = L(e^{j\omega})^*$, and thus in particular

$$|M_i(e^{j\omega})|^2 = |L_i(e^{j\omega})|^2, \quad i = 1, \ldots, q,$$  \hspace{1cm} (7)

From Theorem 2, the sensitivity of the first stage for input signals that are adjacent according to (1) is $k_1\|G\|_2$. Using Theorem 1, adding a white Gaussian noise $w$ to the output of $G$ with covariance matrix $k_1^2\kappa_3^2\|G\|_2^2I_q$ is sufficient to ensure that the signal $v$ on Fig. 1 is differentially private. The MSE for this mechanism can be expressed as

$$\xi(G) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^T \mathbb{E} \left[ \|\left(\frac{F}{L} u_t - (FLG) u_t - (FLw)\|_2 \right)^2 \right]$$

$$\xi(G) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^T \mathbb{E} \left[ \|FLw\|_2 \right] = k_1^2\kappa_3^2\|G\|_2^2\|FL\|_2^2.$$  \hspace{1cm} (8)

We are thus led to consider the minimization of $\|FL\|_2^2\|G\|_2^2$ over the pre-filters $G$. Recall in the following calculation that $|\cdot|$ is also used to denote the Euclidean norm in $\mathbb{C}^d$, for any $d$. We have

$$\|FL\|_2^2\|G\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(L^*(e^{j\omega})F^*(e^{j\omega})F(e^{j\omega})L(e^{j\omega}))d\omega \times$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(G^*(e^{j\omega})G(e^{j\omega}))d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 |M(e^{j\omega})|^2 d\omega \times \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega})|^2 d\omega,$$

where in the last equality we used (6). Now consider on the space of $2\pi$-periodic functions with values in $\mathbb{C}^q$ the inner product $\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{j\omega})^* g(e^{j\omega})d\omega$. By the Cauchy-Schwarz inequality for this inner product applied to the functions $\omega \mapsto |F(e^{j\omega})|^2|M(e^{j\omega})|$ and $\omega \mapsto |G(e^{j\omega})|^2$, we obtain the following bound

$$\|FL\|_2^2\|G\|_2^2 \geq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 \sum_{i=1}^q M_i(e^{j\omega}) G_i(e^{j\omega})d\omega \right)^2$$

i.e., using (7), $\|FL\|_2^2\|G\|_2^2 \geq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|d\omega \right)^2$.

Moreover, the two sides in the Cauchy-Schwarz inequality are equal, i.e., the bound is attained, if $|F(e^{j\omega})|^2|M(e^{j\omega})| = G(e^{j\omega})$. Note that this condition does not depend on $q$. Hence we can simply take $q = 1$, and $L(z) = 1/G(z)$, to get $|F(e^{j\omega})|^2L^*(e^{j\omega}) = G(e^{j\omega})^*$, i.e.,

$$|G(e^{j\omega})|^2 = |F(e^{j\omega})|.$$  \hspace{1cm} (9)

Finding $G$ SISO satisfying (8) is a spectral factorization problem. We can choose $G$ stable and minimum phase, so that its inverse is also stable. The following theorem generalizes [8, Theorem 8].
Fig. 2. (Suboptimal) ZFE mechanism for a MIMO system $F_u = \sum_{i=1}^{m} F_i u_i$, and a diagonal pre-filter $G(z) = \text{diag}(G_{11}(z), \ldots, G_{mn}(z))$. Here $F_i(z)$ is a $p \times 1$ transfer matrix, for $i = 1, \ldots, m$. The signal $w$ is a white Gaussian noise with covariance matrix $\kappa_{\delta,\epsilon}^2\|KG\|^2_{2,m}$.

**Theorem 6:** Let $F$ be a SIMO LTI system with $\|F\|_2 < \infty$. We have, for any LTI system $G$,

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})| d\omega \right)^2,$$

where $|F(e^{j\omega})| = \left( \sum_{i=1}^{p} |F_i(e^{j\omega})|^2 \right)^{1/2}$. If moreover $F$ satisfies the Paley-Wiener condition $\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |F(e^{j\omega})| d\omega > -\infty$, this lower bound on the mean-squared error of the ZFE mechanism can be attained by some minimum phase SISO system $G$ such that $|G(e^{j\omega})|^2 = |F(e^{j\omega})|$, for almost every $\omega \in (-\pi, \pi)$.

**B. MIMO system approximation**

Let us now assume that $F$ has $m > 1$ inputs. We write $F(z) = [F_1(z), \ldots, F_m(z)]$, with $F_i$ a $p \times 1$ transfer matrix. In this case, in view of the complicated expression (5) for the sensitivity of a general MIMO filter, we only provide a suboptimal ZFE mechanism, together with a comparison between the performance of our mechanism and the optimal ZFE mechanism. The idea is to restrict our attention to pre-filters $G$ that are $m \times m$ and diagonal, for which the sensitivity is given in Theorem 3. The problem of optimizing the diagonal pre-filters, using the architecture depicted on Fig. 2, can in fact be seen as designing $m$ SIMO mechanisms.

1) **Diagonal Pre-filter Optimization:** If $G$ is diagonal, then according to Theorem 3 its squared sensitivity is $(\Delta_{2,m}^{m,m} G)^2 = \|KG\|^2_2 = \sum_{i=1}^{m} \|k_i G_{ii}\|^2_2$, with $K = \text{diag}(k_1, \ldots, k_m)$. Following the same reasoning as in the previous subsection, the MSE for this mechanism can be expressed as $\xi(G) = \kappa_{\delta,\epsilon}^2 \|KG\|^2_2 \|FG^{-1}\|^2_2$, with $G^{-1}(z) = \text{diag}(G_{11}(z)^{-1}, \ldots, G_{mn}(z)^{-1})$. Now remark that

$$\|FG^{-1}\|^2_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{m} |F_{i}(e^{j\omega})|^2 |k_i G_{ii}(e^{j\omega})| d\omega.$$ 

Hence from the Cauchy-Schwarz inequality again, we obtain the lower bound

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{m} |F_i(e^{j\omega})|^2 |k_i G_{ii}(e^{j\omega})| d\omega \right)^2.$$ 

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{m} k_i |F_i(e^{j\omega})| d\omega \right)^2,$$

and this bound is attained if $k_i |G_{ii}(e^{j\omega})| = |F_i(e^{j\omega})| |k_i G_{ii}(e^{j\omega})|$, i.e., $k_i |G_{ii}(e^{j\omega})|^2 = |F_i(e^{j\omega})|^2$, $i = 1, \ldots, m$. In other words, the best diagonal pre-filter for the MIMO ZFE mechanism can be obtained from $m$ spectral factorizations of the functions $\omega \mapsto \frac{1}{i\kappa_i} |F_i(e^{j\omega})|$, $i = 1, \ldots, m$.

**Theorem 7:** Let $F = [F_1, \ldots, F_m]$ be a MIMO LTI system with $\|F\|_2 < \infty$. We have, for any diagonal filter $G(z) = \text{diag}(G_{11}(z), \ldots, G_{mn}(z))$,

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{m} k_i |F_i(e^{j\omega})| d\omega \right)^2.$$ 

If moreover each $F_i$ satisfies the Paley-Wiener condition $\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |F_i(e^{j\omega})| d\omega > -\infty$, this lower bound on the mean-squared error of the ZFE mechanism can be attained by some minimum phase systems $G_{ii}$ such that $|G_{ii}(e^{j\omega})|^2 = |F_i(e^{j\omega})|^2$, for almost every $\omega \in (-\pi, \pi)$.

2) **Comparison with Non-Diagonal Pre-filters:** For $F$ a general MIMO system, it is possible that we could achieve better performance with a ZFE mechanism where $G$ is not diagonal, i.e., by combining the inputs before adding the privacy-preserving noise. Here we provide another lower bound on the MSE that one could expect by carrying out this more involved optimization over general pre-filters $G$ rather than just diagonal pre-filters. To simplify the discussion, we assume $k_1 = \ldots = k_m = 1$. The proof of the following result can be found in our report [15].

**Theorem 8:** Consider the same set-up as in Theorem 7, with $k_1 = \ldots = k_m = 1$. The MSE $\xi(G)$ achievable by any $m \times m$ pre-filter $G$ is lower bounded as follows

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \|F(e^{j\omega})\|_{\infty} d\omega \right)^2,$$

where $\|F(e^{j\omega})\|_{\infty} = \text{Tr}(A(e^{j\omega}))$ denotes the nuclear norm of the matrix $F(e^{j\omega})$ (sum of singular values).

The lower bound (11) on the achievable MSE with a general pre-filter in a ZFE mechanism should be compared to the performance (10) that we obtained with diagonal pre-filters (with $k_i = 1$ here). Note that these bounds indeed coincide for $m = 1$.

**V. EXAMPLE: ESTIMATION OF BUILDING OCCUPANCY**

In this section we illustrate some of the ideas discussed above in the context of an application to estimating and forecasting occupancy in an office building equipped with motion detection sensors. As mentioned in Section II-B, such an application raises privacy concerns related to the possibility that some occupants could be tracked individually from the published information, correlated possibly with public information such as the location of their office. The dataset used here comes from a sensor network experiment carried out in the Mitsubishi Electric Research Laboratories (MERL) and described in [17].

The original dataset contains the traces of more than 200 sensors spread over two floors of a building, where each sensor recorded with millisecond accuracy over several months the exact times at which they detected some motion. For
illustration purposes we subsampled the dataset in space and time, summing all the events recorded by several sufficiently close sensors over 5 minute intervals. We formed in this way 10 input signals \( u_i \), \( i = 1, \ldots, 10 \), corresponding to 10 spatial zones (each zone covered by a group of several sensors), with a discrete-time period corresponding to 5 minutes, and \( u_{i,t} \) being the number of events detected by all the sensors in group \( i \) during period \( t \). If we assume say that during a given discrete-time period, a single individual can activate at most 2 sensors in any group, then \( k_i = 2 \) for \( 1 \leq i \leq 25 \). If moreover we assume that any individual travels through at most 5 zones, then we could add a constraint \( \sum_{i=1}^{25} k_i \leq 10 \). Finally we need to assume that a single individual only activates a group of sensor once over the time interval for which we wish to provide differential privacy. Section II-B discussed how to relax this requirement by splitting the input data.

As an illustrative example, we could be interested in a system computing simultaneously and in real-time the following three outputs: i) The sum of the moving averages over the past 30 min for zones 1 to 4; ii) The sum of the moving averages over the past 1 h for zones 3 to 7; iii) A forecast (prediction) of the total number of events detected in the next 20 minutes in all zones, provided by an ARMAX model (with 10 inputs and one output) calibrated using one part of the dataset. Our desired filter has the structure

\[
F = \begin{bmatrix}
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & * & *
\end{bmatrix}, \quad (12)
\]

where * denotes a non-zero transfer function. Fig. 3 shows a sample path over a 25 h period of the 3rd output for a predictive ARMAX model that we designed, as well as a differentially-private-convex version obtained from the diagonal pre-filter optimization procedure of Section IV-B applied to the whole filter \( F \).

VI. CONCLUSION

In this paper we have extended the ZFE mechanism of [7], [8] to the MIMO case. An optimal ZFE mechanism was obtained for the approximation of SIMO filters, and a suboptimal one considering only diagonal pre-filters was obtained for general MIMO filters. Future work includes developing MIMO mechanisms for situations where more information is available about the input signals, e.g., their second-order statistics [9].

REFERENCES