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. In addition, 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 [1], with the goal of vastly improving 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 [2], and that more work is needed to provide rigorous ways to tradeoff privacy with utility.

Privacy-utility tradeoffs can be studied under various definitions of privacy, e.g., k-anonymity [3] or information-theoretic privacy [4], but in recent years, the notion of differential privacy has been particularly popular for applications involving sensitive statistical databases [5], [6]. 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 in this dataset. As a result, the fact for an individual to contribute his data or not does not significantly change the capability of an adversary to make inferences about that individual. One of the main reasons why differentially privacy has proved to be a particularly convenient notion of privacy in practice is that it does not require the modeling of side information to provide guarantees, even though the availability of side information to adversaries is the main reason behind of most privacy breaches [7], [8].

This work was supported by NSERC under Grant RGPIN-435905-13. The authors are with the department of Electrical Engineering, Polytechnique Montreal, QC H3T-1J4, Canada. J. Le Ny is also with GERAD, QC H3T-1J4, Montreal, Canada {jerome.le-ny, meisam.mohammady}@polymtl.ca

In this paper we discuss the application of the notion of differential privacy in the context of real-time system, following our initial work in, e.g., [9]–[11]. In particular, these papers consider a scenario inspired by [12], [13] where a system receives a sensitive input signal recording events related 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 en 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 [14].

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 present certain preliminary calculations necessary to develop our differentially-private dynamic mechanisms, which is the object of Section IV. Finally, in Section V we briefly describe 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 standard abbreviations: SISO for Single-Input Single-Output, SIMO for Single-Input Multiple-Output, and MIMO for Multiple-Input Multiple-Output.

II. PROBLEM STATEMENT

A. Generic Scenario

We consider $m$ sensors detecting events, with sensor $i$ producing a discrete-time scalar signal $\{u_{i,t}\}_{t \geq 0} \in \mathbb{R}$, 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. The privacy constraint is explained and quantified in the next subsection.

**B. Differential Privacy**

As mentioned in the introduction, a differentially private mechanism publishes data in a way that is not too sensitive to the presence or absence of a single individual. A formal definition of differential privacy is provided in Definition [1] below. In the building monitoring application above for 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. Indeed, Wilson and Atkeson [15] for example demonstrate how to track individual users in a building using a network of simple binary sensors such as motion detectors.

1) Adjacency Relation: Formally, we start by defining an symmetric binary relation, denoted $\text{Adj}$, 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, $D := \{ u : \mathbb{N} \rightarrow \mathbb{R}^m \}$, 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

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

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 [10], [12], [13] 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 a differentially private mechanism to offer a guarantee. 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. A first 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. A MIMO mechanism can then treat such data and offer guarantees, addressing one of the main issues for the applicability of the model proposed in [12], [13]. 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: Mechanisms that are differentially private necessarily randomize their outputs, in such a way that they satisfy the following property.

**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 $M : D \times \Omega \rightarrow R$ is $(\epsilon, \delta)$-differentially private for $\text{Adj}$ if for all $d, d' \in D$ such that $\text{Adj}(d, d')$, we have

$$\mathbb{P}(M(d) \in S) \leq e^\epsilon \mathbb{P}(M(d') \in S) + \delta, \forall S \in M. \quad (3)$$

If $\delta = 0$, the mechanism is said to be $\epsilon$-differentially private.

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 $R := \{ y : \mathbb{N} \rightarrow \mathbb{R}^p \}$. 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]. We refer the reader to [9] for a technical discussion on the (standard) $\sigma$-algebra $\mathcal{M}$ used on the output signal space to offer useful guarantees.

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 are interested in evaluating 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

$$\Delta_2^m G = \sup_{\text{Adj}(u,u')} \| Gu - Gu' \|_2 = \sup_{\text{Adj}(u,u')} \| G(u - u') \|_2,$$

where by definition $\| Gu \|_2^2 = \sum_{i=1}^m |G e_i|^2$ and $|x| = (\sum_{k=1}^p |x_k|^2)^{1/2}$ 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 [10]), extending [16], can be used to answer queries in a differentially private way. To present the result, we recall first the definition of the $Q$-function $Q(x) := \int_x^{\infty} e^{-\frac{t^2}{2}} dt$. Now for $\epsilon > 0$, $\delta \geq 0$, let $K = Q^{-1}(\delta)$ and define $\kappa_{\delta, \epsilon} = \frac{1}{2\epsilon}(K + \sqrt{K^2 + 2\epsilon})$.

Theorem 1: Let $G$ be a system with $m$ inputs and $p$ outputs, and with $\ell_2$-sensitivity $\Delta_2 G$ with respect to an adjacency relation $\mathcal{R}$. Then the mechanism $M(u) = Gu + w$, where $w$ is a $p$-dimensional Gaussian white noise with covariance matrix $\kappa_{\delta, \epsilon}^2 (\Delta_2 G)^2 I_p$, is $(\epsilon, \delta)$-differentially private with respect to $\mathcal{R}$. The mechanism $M$ described in Theorem 1, producing a differentially-private version of a system $G$, is called an output-perturbation mechanism. We see that the amount of noise necessary for differential privacy with this mechanism is proportional to the $\ell_2$-sensitivity of the filter and to $\kappa_{\delta, \epsilon}$, which can be shown to behave roughly as $O(\ln(1/\delta))^{1/2}/\epsilon$. Note that we need to add noise proportional to the sensitivity of the whole filter $G$ on each output, even if $G$ was diagonal, say, otherwise trivial attacks that simply average a sufficient number of outputs could potentially detect the presence of an individual with much less overall noise than this output-perturbation scheme, as discussed in Section IV.

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. Moreover, 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. We recall its definition for a system with $m$ inputs

$$||G||_2^2 = \sum_{i=1}^{m} ||G\delta e_i||_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \text{Tr}(G^*(e^{j\omega})G(e^{j\omega}))d\omega.$$  

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

Generalizing the single-input single-output scenario considered in [9], [10] to the case of a system with $m = 1$ but possibly multiple outputs (SIMO), we have immediately the following theorem.

Theorem 2 (SIMO LTI system): Let $G$ be a stable LTI system with one input and $p$ outputs. For the adjacency relation (1),

$$\Delta_2^{1,p} G = k_1||G||_2,$$

where $||G||_2$ is the $H_2$ norm of $G$.

Proof: We have immediately

$$\|G(u - u')\|_2^2 = \|\alpha_1^2\|G\delta_{i_1}||_2^2 \leq k_1^2||G||_2^2,$$

and the bound is attained if $|\alpha_1| = k_1$.

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 $||G||_2^2 = \sum_{i=1}^{m} ||G_{ii}||^2$.

Theorem 3 (Diagonal LTI system): Let $G$ be a stable diagonal LTI system with $m$ inputs and outputs. For the adjacency relation (1),

$$\Delta_2^{m,m} G = ||GK||_2 = \left(\sum_{i=1}^{m} \|k_i G_{ii}\|_2^2\right)^{1/2},$$

where $K = \text{diag}(k_1, \ldots, k_m)$.

Proof: If $G$ is diagonal, then for $u$ and $u'$ adjacent, we have from (2)

$$||G(u - u')||_2^2 = \left\|\sum_{i=1}^{m} \alpha_i G\delta_{i_1} e_i\right\|_2^2 = \left\|\text{col}(\alpha_1 g_{11}^* \delta_{i_1}, \ldots, \alpha_m g_{mm}^* \delta_{i_m})\right\|_2^2,$$

where $\text{col}(x_1, \ldots, x_m)$ denotes a signal with values in $\mathbb{R}^m$ if each $x_i$ is a scalar signal. Here $g_{ii}$ denotes the impulse response of $G_{ii}$. Hence

$$||G(u - u')||_2^2 = \sum_{i=1}^{m} \|\alpha_i g_{ii}^* \delta_{i_1}\|_2^2 = \sum_{i=1}^{m} |\alpha_i|^2 \|G_{ii}\|_2^2,$$

and $|\alpha_i| \leq k_i$ for all $i$. Again 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 $G_i$), such that $||G||_2 < \infty$. For the adjacency relation (1),

$$||GK||_2 \leq \Delta_2^{m,p} G \leq ||G||_2.$$

where $K = \text{diag}(k_1, \ldots, k_m)$ and $|k| = (\sum_{i=1}^{m} k_i^2)^{1/2}$.

Proof: We have

$$G(u - u') = \sum_{i=1}^{m} \alpha_i G\delta_{i_1} e_i,$$
and moreover \( \|G\|_2^2 = \sum_{i=1}^m \|G\delta_i e_i\|_2^2 \) by definition. For the upper bound, we can write
\[
\|G(u-u')\|_2 = \left\| \sum_{i=1}^m \alpha_i G\delta_i e_i \right\|_2 \\
\leq \sum_{i=1}^m |\alpha_i| \|G\delta_i e_i\|_2 \\
\leq |k| \left( \sum_{i=1}^m \|G\delta_i e_i\|_2^2 \right)^{1/2},
\]
where the last inequality results from the Cauchy-Schwarz inequality.

For the lower bound, let us first take \( u' \equiv 0 \). Then consider an adjacent signal with a single discrete impulse on input channel \( i \) of height \( k_i \) at time \( t_i \), \( i = 1, \ldots, m \), with \( t_1 < t_2 < \ldots < t_m \). Let \( \eta > 0 \). Since \( \|G\|_2 < \infty \), \( \|G\delta_i e_i\|_2 < \infty \), and hence \( \|G\delta_i e_i\|_2 \rightarrow 0 \) as \( t \to \infty \). Hence by taking \( t_{i+1} - t_i \) large enough for each \( 1 \leq i \leq m-1 \), i.e., waiting for the effect of impulse \( i \) on the output to be sufficiently small, we can obtain a signal \( u \) such that
\[
\|Gu\|_2^2 = \left\| \sum_{i=1}^m G\delta_i u_i \right\|_2^2 \geq \sum_{i=1}^m k_i^2 \|G\delta_i e_i\|_2^2 - \eta.
\]
Since this is true for any \( \eta > 0 \) and \( \|G\delta_i e_i\|_2^2 = \|G_i\|_2^2 \), we get \( \Delta_2^{m,p}G \geq \|G\|_2^2 \sum_{i=1}^m k_i^2 \|G\|_2^2 \).

Note that if \( k_1 = \ldots = k_m \), the upper bound on the sensitivity is \( k_1\|G\|_2^2 \sqrt{m} \). The squared sensitivity, which is related to the variance of the required amount of noise in an output perturbation scheme, scales then linearly with the number of inputs. We can contrast this bound to the situation where \( G \) is diagonal, in which case the sensitivity is exactly \( k_1\|G\|_2^2 \) from Theorem 3. Moreover, the following example shows that the upper bound of Theorem 4 cannot be improved for the general MISO or MIMO case.

Example 1: Consider the MISO system \( G(z) = [G_{11}(z), \ldots, G_{1m}(z)] \), with \( g_{1i} = \delta_z \), the impulse response of \( G_{1i} \), for some times \( t_1, \ldots, t_m \). Then \( \|G\|_2^2 = m \). Now let \( u' \equiv 0 \) and \( u = \sum_{i=1}^m \delta_z e_i \), so that \( u \) and \( u' \) are adjacent, with \( k_1 = \ldots = k_m = 1 \), and moreover let us choose the times \( t_i \) such that \( t_i + t_i \) is a constant, i.e., take \( t_i = k_i - t_i \) for some \( \kappa \geq \max_i \{t_i\} \). Then \( Gu = \sum_{i=1}^m g_{1i} * u_i = m\delta_e \), and so \( \|Gu\|_2^2 = m^2 \). This shows that the bound of Theorem 4 is tight in this case. Note that this happens because all the events of the signal \( u \) influence the output at the same time. Indeed, if the times \( t_i + t_i \) are all distinct, then we get \( \|Gu\|_2^2 = m \).

C. Exact solution for the MIMO Case

For completeness, we give in this subsection an exact expression for the sensitivity of a MIMO filter. Let \( G \) be a stable LTI system with \( m \) inputs and \( p \) outputs, and state space representation
\[
x_{t+1} = Ax_t + Bu_t, \\
y_t = Cx_t + Du_t,
\]
where the last inequality results from the observability Gramian \( P_0 \), which is the unique positive semi-definite solution of the equation
\[
A^TP_0A - P_0 + C^TC = 0.
\]
Let \( B_t, D_t \) be the \( i \)th column of the matrix \( B \) and \( D \) respectively, for \( i = 1, \ldots, m \). Finally, define for \( i, j \in \{1, \ldots, m\}, i \neq j \), and \( \tau \) in \( Z \)
\[
S_{ij} = \begin{cases} 
B_i^T(A^\tau - 1)^TC^TD_j + B_i^T(A^\tau)^TP_0B_j, & \text{if } \tau > 0 \\
D_i^TB_j + B_i^TP_0B_j, & \text{if } \tau = 0 \\
D_i^TC(A^\tau - 1)B_j + B_i^TP_0A^\tau B_j, & \text{if } \tau < 0.
\end{cases}
\]

Theorem 5: Let \( G \) be a stable LTI system with \( m \) inputs and \( p \) outputs, and state space representation \( 4 \). Then, for the adjacency relation \( 4 \).

\[
\binom{\Delta_2^{m,p}G}{G} \overset{2}{\mathbb{R}}^m = \left\{ G \in \mathbb{R}^{m \times p} : \sup_{t \in \mathbb{N}} \|S_{ij}^{-t} - I\|_2 < \infty \right\}.
\]

Proof: In view of \( 2 \), we have
\[
\Delta_2^{m,p}G = \sup_{|\alpha_i| \leq k_i, t_i > 0} \left\| \sum_{i=1}^m \alpha_i G\delta_i e_i \right\|_2.
\]
For \( y_i = G\delta_i e_i \) and \( y = \sum_{i=1}^m \alpha_i y_i \), we have
\[
\|y\|_2^2 = \sum_{i=0}^\infty \sum_{i=1}^m \alpha_i^2 |y_i|^2 \\
\leq \sum_{i=0}^\infty \sum_{i=1}^m \alpha_i^2 |y_i|^2 + \sum_{i=0}^\infty \sum_{i=1}^m \alpha_i^2 |y_i|^2 = \|y\|_2^2.
\]
where \( K = \text{diag}(k_1, \ldots, k_m) \) and the bound can be attained by taking \( \alpha_i \in \{-k_i, k_i\} \), depending on the sign of \( S_{ij} := \sum_{t=0}^\infty y_i y_j^T \).

Next, we derive the more explicit expression of \( S_{ij} \) given in the theorem. First,
\[
y_{i,t} = \begin{cases} 
0, & t = t_i \\
D_i, & t = t_i \\
CA^{t-t_i}B_i, & t > t_i.
\end{cases}
\]
Assume without loss of generality that \( t_1 \leq t_2 \leq \ldots \leq t_m \). Then if \( t_i = t_j \), we find that
\[
S_{ij} = D_i^TD_j + B_i^TP_0B_j,
\]
with \( P_0 = \sum_{t=0}^\infty (A^t)^TCA^t \) the observability Gramian. If \( t_i < t_j \), then
\[
S_{ij} = B_i^T(A^{t_j-t_i-1})^TC^TD_j + B_i^T(A^{t_j-t_i})^TP_0B_j.
\]
The case \( t_i > t_j \) is symmetric.
D. Discussion

In the expression \( \| \) for the sensitivity, the maximization over inter-event times \( t_i - t_j \) 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.

Still, the expression \( \| \) provides some intuition about the way the system dynamics influence its sensitivity. Note from the expression of \( S_{ij}^T \) in (5) that one way of decreasing the sensitivity of \( G \) is to increase sufficiently the required time \( |t_i - t_j| \) between the events contributed by a single user, in order for \( \| A^{t_i - t_j} \| \) to be small enough. For example, a lower bound on inter-event times in different streams could be enforced in the adjacency relation, which would weaken the differential privacy guarantee but help design mechanisms with better performance. Another possibility would be to have the mechanism ignore events from a given user as long as the lower bound in not reached, provided these events can be easily and immediately identified by the system.

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 [9]. 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 Gaussian white noise signal \( w \) has its variance proportional to the sensitivity of the filter \( G \). It was shown in [9] 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 [17] (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

\[
\sum_{i=1}^{q} L_i(z)G_i(z) = 1.
\]

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

\[
| M_i(e^{j\omega}) |^2 = | L_i(e^{j\omega}) |^2, \quad i = 1, \ldots, q,
\]

and

\[
\sum_{i=1}^{q} M_i(e^{j\omega})^*G_i(e^{j\omega}) = 1.
\]

From Theorem 2, the sensitivity of the first stage for input signals that are adjacent according to (4) is \( k_1 ||G||_2 \). Using Theorem 1, adding a white Gaussian noise \( w \) to the output of \( G \) with covariance matrix \( k_1^2 k_2^2 ||G||_2^2 L_q \) is sufficient to ensure that the signal \( v \) on Fig. [1] is differentially private. The MSE for this mechanism can be expressed as

\[
\xi_1(G) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{\infty} \mathbb{E} \left[ \| (Fu)_{t} - (FLGu)_{t} - (FLw)_{t} \|_{2}^2 \right]
\]

\[
\xi_2(G) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{\infty} \mathbb{E} \left[ \| (FLw)_{t} \|_{2}^2 \right]
\]

\[
\xi_3(G) = k_1^2 k_2^2 ||G||_2^2 \|FL\|_2^2.
\]

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

\[
\|FL\|_2^2 ||G||_2^2
\]

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

\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(G^*(e^{j\omega})G(e^{j\omega}))d\omega
\]

\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 |L(e^{j\omega})|^2 d\omega \times \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega})|^2 d\omega
\]

\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 |M(e^{j\omega})|^2 d\omega \times \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega})|^2 d\omega,
\]

where in the last equality we used (7). Now consider on the space of 2\pi-periodic functions with values in \( \mathbb{C}^q \) the inner product

\[
(f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{j\omega})^* g(e^{j\omega}) d\omega.
\]
By the Cauchy-Schwarz inequality for this inner product applied to the functions \( \omega \mapsto |F(e^{j\omega})|M(e^{j\omega}) \) and \( \omega \mapsto G(e^{j\omega}) \), we obtain the following bound

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

i.e., using (8),

\[
\|FL\|_2^2 \|G\|_2^2 \geq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})| d\omega \right)^2.
\]

Moreover, the two sides in the Cauchy-Schwarz inequality are equal, i.e., the bound is attained, if

\[
|F(e^{j\omega})|M(e^{j\omega}) = G(e^{j\omega}).
\]

Note that this condition does not depend on \( q \). Hence we can simply take \( q = 1 \), and \( L(z) = 1/G(z) \), to get

\[
|F(e^{j\omega})|L^* (e^{j\omega}) = G(e^{j\omega})
\]

i.e., \( |G(e^{j\omega})|^2 = |F(e^{j\omega})| \).

Finding \( G \) SISO satisfying (9) is a spectral factorization problem. We can choose \( G \) stable and minimum phase, so that its inverse is also stable. The following theorem summarizes the preceding discussion and generalizes [10, Theorem 8].

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

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

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 (9) 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. 1, 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

\[
(\Delta_{m,m}^n G)^2 = \|KG\|_2^2 = \sum_{i=1}^{m} \|k_i G_i\|_2^2, \quad \text{with} \quad K = \text{diag}(k_1, \ldots, k_m).
\]

Following the same reasoning as in the previous subsection, the MSE for this mechanism can be expressed as

\[
\xi(G) = k^2_{\delta,\epsilon} \|KG\|_2^2 \|FG^{-1}\|_2^2,
\]

with \( G^{-1}(z) = \text{diag}(G_{11}(z)^{-1}, \ldots, G_{mm}(z)^{-1}) \). Now remark that

\[
\|FG^{-1}\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{m} \|F_i(e^{j\omega})\|^2 G_i(e^{j\omega})^2 d\omega.
\]

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

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

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

and this bound is attained if

\[
k_i |G_i(e^{j\omega})| = |F_i(e^{j\omega})| \quad \text{for each} \quad i = 1, \ldots, m.
\]

In other words, the best diagonal pre-filter for the MIMO ZFE mechanism can be obtained from \( m \) spectral factorizations of the functions \( \omega \mapsto \frac{1}{k_i} |F_i(e^{j\omega})| \), \( i = 1, \ldots, m \).

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

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

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 \( |G_{ii}(e^{j\omega})|^2 = |F_i(e^{j\omega})| \), for almost every \( \omega \in [-\pi, \pi] \).

![Diagram](image-url)
2) Comparison with Non-Diagonal Pre-filters: For 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 $k_1 = \ldots = k_m = 1$.

Hence consider a general $m \times m$ pre-filter $G$ with left inverse $L$. With the lower bound of Theorem 4 designing a ZFE mechanism based on sensitivity as above would require adding a noise with variance at least $\kappa_3^2 \|G\|_2^2$. This would lead to an MSE at least equal to $\kappa_3^2 \|G\|_2^2 \|FL\|_2^2$. Now note that

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

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

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(A(e^{j\omega})^2L(e^{j\omega})L(e^{j\omega})^*)d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(A(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^*)A(e^{j\omega})d\omega,$$

where for all $\omega$, $A(e^{j\omega})$ is the positive-semidefinite square root of $F(e^{j\omega})^*F(e^{j\omega})$, i.e., $A(e^{j\omega})^2 = F(e^{j\omega})^*F(e^{j\omega})$. Then, once again from the Cauchy-Schwarz inequality

$$\|G\|_2^2 \|FL\|_2^2 = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(G(e^{j\omega})^*G(e^{j\omega}))d\omega \right]$$

$$\times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(A(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^*)A(e^{j\omega})d\omega \right]$$

$$\|G\|_2^2 \|FL\|_2^2 \geq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Tr}(A(e^{j\omega})L(e^{j\omega})G(e^{j\omega})d\omega \right)^2$$

$$\xi(G) \geq \kappa_3^2 \epsilon \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \|F(e^{j\omega})\|_*d\omega \right)^2,$$

(12)

where $\|F(e^{j\omega})\|_* = \text{Tr}(A(e^{j\omega}))$ denotes the nuclear norm of the matrix $F(e^{j\omega})$ (sum of singular values). The lower bound (12) on the achievable MSE with a general pre-filter in a ZFE mechanism should be compared to the performance (11) 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 [18] and on Fig. 3.

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 $u_{i,t}$, $i = 1, \ldots, 10$, 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 $\sum_i k_i \leq 10$. 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:

- The sum of the moving averages over the past 30 min for zones 1 to 4.
- The sum of the moving averages over the past 1 h for zones 3 to 7.
- A forecast (prediction) of the total number of events detected in the next 20 minutes in all zones, provided by an ARMAX model [19] (with 10 inputs and one output) calibrated using one part of the dataset.

Hence our desired filter has the structure

$$F = \begin{bmatrix} * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & * \end{bmatrix},$$

(13)

where * denotes a non-zero transfer function. Fig. 4 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 version obtained from the diagonal pre-filter optimization procedure of Section IV-B applied to the whole filter $F$. Notice on the figure that the approach used here relying on the notion of sensitivity requires a noise level independent of the size of the desired output signal, hence low signal values tend to be easily buried in the noise.

VI. CONCLUSION

In this paper we have extended the ZFE mechanism of [9], [10] 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, in which case one can improve on the ZFE mechanism [11].

REFERENCES

[1] E. A. Lee, J. D. Kubiatowicz, J. M. Rabaey, A. L. Sangiovanni-Vincentelli, S. A. Seshia, J. Wawrzynek, D. Blauw, P. Dutta, K. Fu, C. Guestrin, R. Jafare, D. Jones, V. Kumar, and R. Murray, “The Terrawarm Research Center,” University of California at Berkeley, Tech. Rep., 2012.

[2] H. Chan and A. Perrig, “Security and privacy in sensor networks,” Computer, vol. 36, no. 10, pp. 103–105, Oct 2003.

[3] L. Sweeney, “k-anonymity: A model for protecting privacy,” International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 05, pp. 557–570, 2002.

[4] L. Sankar, S. R. Rajagopalan, and H. V. Poor, “A theory of privacy and utility in databases,” Princeton University, Tech. Rep., February 2011.

[5] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in Proceedings of the Third Theory of Cryptography Conference, 2006, pp. 265–284.

[6] C. Dwork, “Differential privacy,” in Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP), ser. Lecture Notes in Computer Science, vol. 4052, Springer-Verlag, 2006.

[7] A. Narayanan and V. Shmatikov, “Robust de-anonymization of large sparse datasets (how to break anonymity of the Netflix Prize dataset),” in Proceedings of the IEEE Symposium on Security and Privacy, 2008.

[8] J. A. Calandrino, A. Kilzer, A. Narayanan, E. W. Felten, and V. Shmatikov, “you might also like”: Privacy risks of collaborative filtering,” in Proceedings of the IEEE Symposium on Security and Privacy, Berkeley, CA, May 2011.

[9] J. Le Ny and G. J. Pappas, “Differentially private filtering,” in Proceedings of the Conference on Decision and Control, Maui, HI, December 2012.

[10] ——, “Differentially private filtering,” IEEE Transactions on Automatic Control, vol. 59, no. 2, pp. 341–354, February 2014.

[11] J. Le Ny, “On differentially private filtering for event streams,” in Proceedings of the Conference on Decision and Control, Florence, Italy, December 2013.

[12] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, “Differential privacy under continual observations,” in Proceedings of the ACM Symposium on the Theory of Computing (STOC), Cambridge, MA, June 2010.

[13] T.-H. H. Chan, E. Shi, and D. Song, “Private and continual release of statistics,” ACM Transactions on Information and System Security, vol. 14, no. 3, pp. 26:1–26:24, November 2011.

[14] J. Le Ny, A. Touati, and G. J. Pappas, “Real-time privacy-preserving model-based estimation of traffic flows,” in Proceedings of the Fifth International Conference on Cyber-Physical Systems (ICCPSS), April 2014.

[15] D. H. Wilson and C. Atkeson, “Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors,” in Persisve Computing, ser. Lecture Notes in Computer Science, H.-W. Gellersen, R. Want, and A. Schmidt, Eds. Springer Berlin Heidelberg, 2005, vol. 3468, pp. 62–79.

[16] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, “Our data, ourselves: Privacy via distributed noise generation,” Advances in Cryptology-EUROCRYPT 2006, pp. 486–503, 2006.

[17] J. Cao, Q. Xiao, G. Ghinita, N. Li, E. Bertino, and K.-L. Tan, “Efficient and accurate strategies for differentially-private sliding window queries,” in Proceedings of the International Conference on Extending Database Technology, 2013.

[18] C. Wren, Y. Ivanov, D. Leigh, and J. Westhues, “The MERL motion detector dataset,” Mitsubishi Electric Research Laboratories, Tech. Rep. TR2007-069, November 2007.

[19] L. Ljung, System Identification: Theory for the User, ser. Information and System Sciences. Prentice Hall, 1998.