Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 52-20 (2019) 303–308

Differentially Private Event-Triggered Sampling ⋆

Jerome Le Ny * **Sandra Hirche** **

* *Polytechnique Montreal and GERAD, Montreal, QC H3T1J4 Canada (e-mail: jerome.le-ny@polymtl.ca)*

** *Chair for Information-oriented Control, Technical University of Munich, Arcisstraße 21, D-80290 Munich, Germany (e-mail: hirche@tum.edu)*

Abstract: This paper describes a differentially private event-triggered sampling mechanism to select measurement samples from a data sequence whose dynamics can be modelled by a stochastic linear system. The mechanism produces subsequences that can be used to reestimate the original sequence relatively accurately and the differential privacy constraint guarantees that these subsequences are insensitive to certain variations in the input sequence. The subsampling process can be motivated by the presence of communication bandwidth constraints, but also provides an additional tool to explore achievable privacy-utility tradeoffs in privacy-preserving signal processing and control. Event-triggered sampling can offer benefits over periodic subsampling by attempting to select the most useful samples, but the fact that it leaks information when no sampling occurs must be carefully taken into account to meet the differential privacy requirement. We propose a design using a stochastic sampling threshold, leveraging the "sparse vector technique" from differential privacy to incur a privacy loss only when samples are actually released. This design includes a suboptimal but tractable recursive finite-dimensional estimator that can also be used to re-estimate the original sequence from the differentially private noisy subsequence.

Keywords: Data privacy, Event-triggered sampling, Sampling systems, State estimation, Recursive estimation.

1. INTRODUCTION

Because many initiatives to build a more intelligent infrastructure rely on the collection and processing of a large amount of personal information, often continuously and in real time, there is a growing interest in being able to better measure and control the impact of large-scale automated systems such as intelligent transportation systems or smart grids on the privacy of individuals (President's Council of Advisors on Science and Technology, 2014). Among various definitions of privacy that have been proposed in the past two decades for privacy-preserving data release and analysis, differential privacy (Dwork et al., 2006; Dwork and Roth, 2014) is one of the most successful notions and has been applied to the analysis of datasets ranging from data collected via web browsers (Erlingsson et al., 2014) to the U.S. Census Data (Abowd, 2018). Differentially private algorithms publish statistics about a protected dataset that are insensitive to specific variations in the dataset (e.g., adding or removing a single person's data), to ensure that certain inferences about this dataset become provably difficult (e.g., detecting if someone's data is present in the dataset). These inferences could be based on the published statistics and any source of side information.

For the differentially private real-time processing of signals, previous work has focused on signal perturbation schemes using additive noise, see, e.g., Le Ny and Pappas (2014). The potential benefits of first selecting a subset of a sensitive signal's samples, especially when a model is available to estimate and predict this signal, are comparatively unexplored except in Fan and Xiong (2014), although (random) subsampling is known to enhance differential privacy guarantees for the analysis of static datasets (Li et al., 2012). This motivates the problem considered in this paper, which is to select in real-time which samples of a sensitive data stream to release to simultaneously meet a differential privacy requirement and maintain the ability to reestimate the initial data stream relatively accurately. In addition to offering an additional tool to explore privacy-accuracy tradeoffs, subsampling might also be necessary in the context of networked monitoring and control systems to meet certain communication bandwidth constraints for example.

We focus here on an event-triggered sampling approach (Heemels et al., 2012), where a data sample is only released when it deviates sufficiently from a value that can be predicted based on previously released samples together with a dynamic model assumed to be publicly available. Our work is motivated by the following idea: in situations where publishing more samples of a signal in a differentially private way requires more noise, the overall impact on accuracy of having one more sample might

⋆ This work was performed while the first author was on sabbatical leave at the Technical University of Munich, and supported by NSERC under Grants RGPAS-522686-2018 and RGPIN-5287-2018, and by a fellowship from the Alexander von Humboldt Foundation.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2019.12.172

become negative due to the additional noise, especially if the sample can be predicted accurately by the model.

The paper is organized as follows. Section 2 states the problem of designing a differentially private event-triggered sampling mechanism and recalls some basic facts about differential privacy. The proposed mechanism is described in Section 3, together with a proof of the differential privacy guarantees it provides. Finally, Section 4 focuses on the design of an estimator embedded in the sampling mechanism to compute the triggering condition at each period.

Notation. The Laplace distribution with mean zero and variance $2/\lambda^2$ is denoted $\text{Lap}(\lambda)$ and has probability density function (pdf) $\frac{\lambda}{2}\exp(-\lambda|x|)$ for $x \in \mathbb{R}$. The exponential distribution with mean $1/\lambda$ and variance $1/\lambda^2$ is denoted $\exp(\lambda)$ and has pdf $\lambda\exp(-\lambda x)$ for $x \geq 0$. We use the notation $\|M\|_1$ for the induced 1-norm of an $m \times n$ matrix $M = [m_{ij}]_{i,j}$, i.e., $\|M\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^{m}|m_{ij}|$. For M symmetric positive definite, $M^{-1/2}$ denotes the inverse of the principal square root of M. A finite sequence $\{x_k\}_{s\leq k\leq t}$ is denoted $x_{s:t}$.

2. PROBLEM STATEMENT

Consider a private signal $x := \{x_k\}_{k\geq 0}$, with $x_k \in \mathbb{R}^n$, captured over time by a data aggregator. A mathematical model taking the form of a stochastic linear time-invariant system is publicly known for this observed signal

$$x_{k+1} = Ax_k + w_k,\ k \geq 0, \tag{1}$$

where the initial state $x_0 \in \mathbb{R}^n$ is assumed to have mean $\bar{x}_0$ and covariance matrix $\bar{\Sigma}_0$, the matrix $A \in \mathbb{R}^{n\times n}$ is known, and w_k is a zero-mean white noise signal with known covariance matrix W. The data aggregator would like to release publicly a (perturbed) *subsequence* of the whole sequence $\{x_k\}_{k\geq 0}$. More precisely, it needs to decide at each period k if the sample x_k should be released or not. This could be due for example to the necessity of transmitting the samples over a communication network subject to a bandwidth constraint. In addition, it is required that the released subsequence be differentially private, as defined in the next subsection. In that regard, subsampling might also be beneficial to meet the privacy constraint, independently of any other communication bandwidth constraint.

2.1 Differential Privacy

In this section we define formally the differential privacy requirement for the published subsequence. A mechanism publishing the result of a computation performed on a protected dataset is differentially private (Dwork et al., 2006) if: i) it randomizes its answer and ii) its answer's probability distribution does not depend too strongly on certain specified variations in the dataset. These variations can be interpreted as the impact that a single individual's data has on the dataset. More precisely, given a space $\mathcal{D}$ of datasets, we start by defining a binary symmetric relation on $\mathcal{D}$, called adjacency and denoted Adj, which captures which varations we allow. A typical example would be to define adjacent datasets as those for which the data of exactly one individual has been added or removed. For an algorithm (also called mechanism) to be differential private, it should be hard for a third party to decide, based on observing only the output of the mechanism, if the dataset used was d or d', for any two adjacent d and d'. For this purpose, denoting M the mechanism, $M(d)$ is a random variable and differential privacy asks that the distributions $M(d)$ and $M(d')$ for any two adjacent datasets d and d' be close, as expressed in (2) below.

Definition 1. Let $\mathcal{D}$ be a space equipped with a given symmetric binary relation Adj, and $(\mathcal{R}, \mathcal{M})$ be a measurable space, where $\mathcal{M}$ is a given σ-algebra over $\mathcal{R}$. Let $\epsilon \geq 0$. A randomized mechanism M from $\mathcal{D}$ to $\mathcal{R}$ is ϵ-differentially private (for Adj) if for all $d, d' \in \mathcal{D}$ such that $(d, d') \in \text{Adj}$, for all sets S in $\mathcal{M}$,

$$\mathbb{P}(M(d) \in S) \leq e^{\epsilon}\,\mathbb{P}(M(d') \in S). \tag{2}$$

In this paper, we take $\mathcal{D} = (\mathbb{R}^n)^{\mathbb{N}}$ and we consider that two data sequences x, x' in $\mathcal{D}$ are adjacent if an only if

$$\|x_k - x'_k\|_1 \leq \rho,\ \ \forall k \geq 0. \tag{3}$$

In other words, we allow at each period a variation of at most ρ in 1-norm for the sample x_k. For example, x might represents counts reported by by n motion detection sensors and a single individual is assumed to activate at most ρ of these sensors at each period to satisfy (3). See Dwork et al. (2010) or Le Ny and Mohammady (2018) for a discussion of some application examples.

We now informally state a few required results. First, an important property of differential privacy is its "resilience to post-processing", which says that further transforming the output $M(d)$ of a differentially private mechanism, without re-accessing the dataset d, cannot weaken the differential privacy guarantee provided by M (Dwork and Roth, 2014), (Le Ny and Pappas, 2014, Theorem 1). Hence, post-processing a released subsequence of samples by an estimator does not change the privacy guarantee. Second, we can compose differentially private mechanisms, with the following degradation of the privacy parameter. If we have N mechanisms $M_1 : \mathcal{D} \to \mathcal{R}_1$, $M_k : \mathcal{D} \times \mathcal{R}_1 \times \ldots \times \mathcal{R}_{k-1} \to \mathcal{R}_k$, $2 \leq k \leq N$, and each M_i is ϵ_i-diffentially privacy with respect to its first argument for each fixed value of its subsequent arguments, then the composed mechanism releasing $(d_1, \ldots, d_N)$ with each output defined recursively as $d_1 = M_1(d)$ and $d_k = M_k(d, d_1, \ldots, d_{k-1})$ for $k \geq 2$, is $(\epsilon_1 + \ldots + \epsilon_N)$-differentially private (Dwork and Roth, 2014, Corollary B.2). Finally, the sensitivity (also called 1-sensitivity) of a function $q : \mathcal{D} \to \mathbb{R}^n$ is defined as $\Delta q := \sup_{d,d':\text{Adj}(d,d')} \|q(d) - q(d')\|_1$. Then, the Laplace mechanism (Dwork et al., 2006) $M : \mathcal{D} \to \mathbb{R}^n$ defined as $M(d) = q(d) + \xi$, with $\xi \in \mathbb{R}^n$ a random vector with iid components $\xi_i \sim \text{Lap}(\epsilon/\Delta q)$, is ϵ-differentially private. Note that this noise has variance $2\Delta q^2/\epsilon^2$ on each component.

2.2 Event-Triggered Sampling Approach

We are concerned in this paper with the problem of selecting which samples of the sequence x to release, while guaranteeing that the released subsequence be differentially private for the adjacency relation (3). Consider the following three possible mechanisms.

(i) Subsample the signal x periodically and perturb the selected samples using to the Laplace mechanism.

(ii) Perturb the whole sequence x using the Laplace mechanism, then subsample that perturbed sequence using an event-triggered mechanism.
(iii) Subsample x using an event-triggered mechanism, and perturb the selected samples using to the Laplace mechanism.

Scheme (i) is the most straightfoward to implement, but as with any periodic sampling scheme, not considering the value of the samples in their selection process might result in suboptimal performance for networked monitoring and control systems (Heemels et al., 2012). Scheme (ii) has the important drawback that the privacy-preserving noise is introduced before sampling, hence the Laplace mechanism will result in a degradation of the privacy parameter proportional to the length of the whole sequence, not the number of ultimately selected samples. An alternative is explored by Fan and Xiong (2014), where a self-triggering scheme is proposed to adapt the sampling rate to the size of the innovations at the sampling times only. Scheme (iii) is the focus of this paper. The difficulty is to be able to subsample based on considering the values of the samples, yet avoid the same worst-case privacy degradation as with scheme (ii). A direct application of the composition mechanism recalled above would give a degradation proportional to the length of the whole sequence. Our goal instead is to obtain an event-triggered mechanism with privacy degradation proportional to the number of selected samples.

3. DIFFERENTIALLY PRIVATE EVENT-TRIGGERED SAMPLER

The architecture of the proposed event-triggered sampler is shown on Fig. 1. A sample x_t is first selected or rejected by the event detector based on a condition involving the size of the difference $x_k - \hat{x}_k$ between the value of the sample and a value predicted based on past selected samples, after a sanitization process involving additive noise ξ to enforce differential privacy. Let $\gamma_k \in \{\top, \bot\}$ be a variable representing the output of the event detector, with $\gamma_k = \top$ if a sample x_k is selected at period k, and $\gamma_k = \bot$ otherwise. The subsampled sequence $\tilde{x}^s_k$ on the figure is defined as

$$\tilde{x}^s_k = \begin{cases} x_k + \xi_k, & \text{if } \gamma_k = \top \\ \bot, & \text{if } \gamma_k = \bot. \end{cases}$$

In particular, if no sampling occurs at period k, then this information is passed to the subsequent estimator. If sampling occurs at period k, then a random variable ξ_k is added to the sample. The estimator produces an estimate $\hat{x}$ based on the sequence $\tilde{x}^s$, which therefore includes the past selected and perturbed samples, as well as the information about the periods at which no sampling occurred. The signals $\tilde{x}^s$, $\hat{x}$, γ (or a subset of these) can be published by the mechanism, as these signals are differentially private. In particular, since our estimator is suboptimal, it can be useful to output the noisy samples $\tilde{x}^s_k$, if another estimator with possibly better performance is to be implemented outside of the proposed sampling scheme. In a remote estimation scenario with bandwidth constraints, as in Wu et al. (2016) for example, one can also transmit just the noisy samples and reproduce the estimator at the remote location. Note that by the resilience to post-processing property, any such subsequent estimator still produces differentially private outputs.

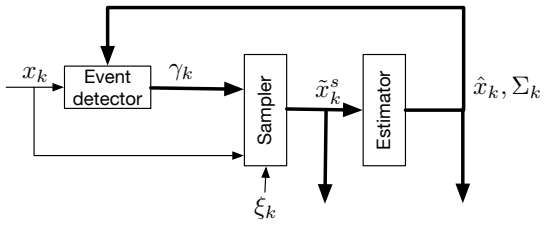

Fig. 1. Structure of the differentially private event-based sampler. The signals $\gamma, \tilde{x}^s, \hat{x}$ represented by thick lines are differentially private and therefore can be published.

The rest of this section is dedicated to describing more precisely the sample selection method (event detector) and to proving the differential privacy property of the sequences γ, $\tilde{x}^s$, $\hat{x}$. Section 4 describes in more details our estimator design.

3.1 Detecting the First Event

The event detector relies on a stochastic thresholding policy. This is reminiscent of recent work discussing how stochastic thresholding simplifies the design of subsequent Bayesian state estimators for certain linear systems driven by Gaussian noise (Han et al., 2015; Wu et al., 2016). However, the main motivation in this work for using a stochastic thresholding policy is for the purpose of guaranteeing differential privacy with a privacy degradation that depends only on the number of ultimately selected samples, following the "sparse vector technique" idea from the differential privacy literature (Dwork et al., 2009; Dwork and Roth, 2014; Lyu et al., 2017),

In order to present the formal privacy results, we first describe how the event detector operates from period $k = 0$ until the first time a sample is selected, assuming for now that the detector stops releasing outputs at that time. Let $\lambda_\tau, \lambda_\nu$ be two positive parameters, with $\lambda_\tau \neq \lambda_\nu$. Initially, we set $z_0 = x_0 - \bar{x}_0$ and we generate a random variable $\tau \sim \exp(\lambda_\tau)$, whose value is not published. Note that the covariance matrix of z_0 is provided by the model, equal to $\bar{\Sigma}_0$. Then, at each subsequent period $k \geq 0$, given some values for $z_k, \bar{\Sigma}_k$, we generate a fresh random variable $\nu_k \sim \text{Lap}(\lambda_\nu)$ and output the following decision

$$\begin{cases} \gamma_k = \bot \text{ (idle)}, & \text{if } \nu_k \geq f_k(z_k) - \tau, \\ \gamma_k = \top \text{ (event detected, stop)}, & \text{otherwise}, \end{cases} \tag{4}$$

where

$$f_k(z_k) = \frac{1}{\|\bar{\Sigma}_k^{-1/2}\|_1} \|\bar{\Sigma}_k^{-1/2} z_k\|_1.$$

If $\gamma_k = \bot$, we update the parameters for the next period, as follows

$$\bar{x}_{k+1} = A\bar{x}_k \;\; (= A^k \bar{x}_0) \tag{5}$$

$$z_{k+1} = x_{k+1} - \bar{x}_{k+1} \tag{6}$$

$$\bar{\Sigma}_{k+1} = \eta_k(\lambda_\tau, \lambda_\nu) A \bar{\Sigma}_k A^T + W, \tag{7}$$

where $\bar{x}_k$ is a state estimate and $\eta_k(\lambda_\tau, \lambda_\nu) \in \mathbb{R}_+$ is a positive scalar coefficient, strictly smaller than 1, which is computed in Section 4. Note that the iterations (5) and (7) do not depend on the values x_k of the data sequence.

Denote by $k_s \geq 0$ the time period at which the event detector stops, i.e., $\gamma_{k_s} = \top$. The following proposition is an immediate consequence of the triangle inequality.

Proposition 2. The sensitivity of f_k is bounded above by ρ, for all $0 \leq k \leq k_s$.

Proof. Define $z_k = x_k - \bar{x}_k$ and $z'_k = x'_k - \bar{x}_k$, for x, x' adjacent sequences according to (3) and $\bar{x}_k$ defined from $\bar{x}_0$ and the iterations (5). Then

$$\|z_k - z'_k\|_1 = \|x_k - x'_k\|_1 \leq \rho.$$

Hence, by the triangle inequality,

$$\begin{aligned} |f_k(z_k) - f_k(z'_k)| &\leq \frac{1}{\|\Sigma_k^{-1/2}\|_1} \|\Sigma_k^{-1/2}(z_k - z'_k)\|_1 \\ &\leq \frac{\|\Sigma_k^{-1/2}\|_1}{\|\Sigma_k^{-1/2}\|_1} \|z_k - z'_k\|_1 \leq \rho. \end{aligned}$$

Theorem 3. If the first event time k_s is finite, the event detector described above, publishing the sequence $\gamma_{0:k_s}$, is ϵ-differentially private, with $\epsilon = \rho(\lambda_\tau + 2\lambda_\nu)$.

Remark 4. Compared to other event-triggered sampling schemes described in the networked control literature (see, e.g., Wu et al. (2016)), the introduction of the random variable τ (which is not updated with k), as well as the choice of the Laplace distribution for ν_k, are crucial here to obtain a differential privacy guarantee. The sparse vector technique normally uses also the Laplace distribution for τ (Lyu et al., 2017), but the exponential distribution we use here is also appropriate and leads to a slightly simpler estimator design in Section 4.

Proof. The proof is based on a variation of the sparse vector technique from the differential privacy literature. Consider a sequence $z = \{z_k\}_{k\geq 0}$. For this sequence let $\bar{\gamma} = \{\bar{\gamma}_0, \ldots, \bar{\gamma}_{k_s}\}$ be a sequence of decisions (4) produced by the stochastic threshold detector, which is necessarily of the form $\bar{\gamma}_0 = \perp, \ldots \bar{\gamma}_{k-1} = \perp, \bar{\gamma}_{k_s} = \top$. Now, for $0 \leq k \leq k_s$, let $A_k(z)$ denote the event

$$A_k(z_k) = \{\nu_k \geq f_k(z_k) - \tau\},$$

and define, for any $\alpha \geq 0$, the conditional probabilities

$$\begin{aligned} g_k(\alpha, z) &= \mathbb{P}(\nu_k \geq f_k(z) - \tau | \tau = \alpha) \\ g_k^c(\alpha, z) &= \mathbb{P}(\nu_k < f_k(z) - \tau | \tau = \alpha). \end{aligned}$$

Then, we have

$$\begin{aligned} \mathbb{P}(\gamma(z) = \bar{\gamma}) &= \mathbb{P}\left(\cap_{k=0}^{k_s-1} A_k(z_k) \cap A_{k_s}^c(z_{k_s})\right) \\ &= \int_{\alpha=0}^{\infty} \left[\prod_{k=0}^{k_s-1} g_k(\alpha, z_k)\right] g_{k_s}^c(\alpha, z_{k_s})\, p_\tau(\alpha) d\alpha, \end{aligned}$$

where $p_\tau(\alpha) = \lambda_\tau e^{-\lambda_\tau \alpha}$ is the probability density of τ. Now, for z' adjacent to z, we have for all k

$$f_k(z'_k) - \Delta \leq f_k(z_k) \leq f_k(z'_k) + \Delta.$$

with $\Delta = \rho$ by Proposition 2. Hence,

$$\begin{aligned} &\mathbb{P}(\nu_k \geq f_k(z_k) - \alpha) \leq \mathbb{P}(\nu_k \geq f_k(z'_k) - \alpha - \Delta) \\ &\text{i.e., } g_k(\alpha, z_k) \leq g_k(\alpha + \Delta, z'_k). \end{aligned}$$

Similarly,

$$\begin{aligned} &\mathbb{P}(\nu_k < f_k(z_k) - \alpha) \leq \mathbb{P}(\nu_k < f_k(z'_k) - \alpha + \Delta) \\ &\text{i.e., } g_k^c(\alpha, z_k) \leq g_k^c(\alpha - \Delta, z'_k). \end{aligned}$$

We then have

$$\begin{aligned} &\mathbb{P}(\gamma(z) = \bar{\gamma}) \\ &\leq \int_{\alpha=0}^{\infty} \left[\prod_{k=0}^{k_s-1} g_k(\alpha + \Delta, z'_k)\right] g_{k_s}^c(\alpha - \Delta, z'_k)\, p_\tau(\alpha) d\alpha, \\ &\leq \int_{\beta=\Delta}^{\infty} \left[\prod_{k=0}^{k_s-1} g_k(\beta, z'_k)\right] g_{k_s}^c(\beta - 2\Delta, z'_{k_s})\, p_\tau(\beta - \Delta) d\beta, \\ &\leq \int_{\beta=0}^{\infty} \left[\prod_{k=0}^{k_s-1} g_k(\beta, z'_k)\right] g_{k_s}^c(\beta - 2\Delta, z'_{k_s})\, p_\tau(\beta - \Delta) d\beta, \end{aligned}$$

where we have used the change of variable $\beta = \alpha + \Delta$ and the fact that the integrand is nonnegative. Now, we have

$$p_\tau(\beta - \Delta) = \lambda_\tau e^{-\lambda_\tau(\beta - \Delta)} \leq e^{\lambda_\tau \Delta} p_\tau(\beta).$$

In addition,

$$\begin{aligned} g_{k_s}^c(\beta - 2\Delta, z'_{k_s}) &= \mathbb{P}(\nu_{k_s} < f_{k_s}(z'_{k_s}) - \beta + 2\Delta) \\ &= \frac{\lambda_\nu}{2} \int_{-\infty}^{f_{k_s}(z'_{k_s}) - \beta + 2\Delta} e^{-\lambda_\nu |v|} dv \\ &= \frac{\lambda_\nu}{2} \int_{-\infty}^{f_{k_s}(z'_{k_s}) - \beta} e^{-\lambda_\nu |u + 2\Delta|} du \\ &\leq e^{2\lambda_\nu \Delta} g_{k_s}^c(\beta, z'_{k_s}), \end{aligned}$$

using the change of variable $u = v - 2\Delta$ and the fact that $|u + 2\Delta| \geq |u| - 2\Delta$ by the triangle inequality. In conclusion, we finally get

$$\mathbb{P}(\gamma(z) = \bar{\gamma}) \leq e^{\Delta(\lambda_\tau + 2\lambda_\nu)}\, \mathbb{P}(\gamma(z') = \bar{\gamma}). \qquad \square$$

3.2 Releasing Samples and Iterating

Suppose now that at the first period k_s where the event detector of the previous section produces $\gamma_{k_s} = \top$, we publish the corresponding signal sample perturbed by Laplace noise

$$\tilde{x}_{k_s}^s = x_{k_s} + \xi_{k_s}, \;\; \xi_{k_s} \sim \text{Lap}(\lambda_x).$$

Then, by Theorem 3 and the basic composition theorem for differential privacy, the mechanism releasing this first sample is ϵ-differentially private, with $\epsilon = \rho(\lambda_\tau + 2\lambda_\nu + \lambda_x)$.

We can now complete the description of the event-triggered sampling mechanism of Fig. 1, by describing the recursive algorithm for the estimator block. We initialize $\bar{x}_0$ and $\bar{\Sigma}_0$ at the values provided by the model. Then, at each period $k \geq 0$, the algorithm reads $\tilde{x}_k^s$. If $\tilde{x}_k^s = \perp$, it outputs $\hat{x}_k = \bar{x}_k, \Sigma_k = \eta_k(\lambda_\tau, \lambda_\nu)\bar{\Sigma}_k$, with η_k computed in Section 4. If instead $\tilde{x}_k^s$ is a true noisy sample, then it outputs the result of a measurement update step

$$\hat{x}_k = \bar{x}_k + \bar{\Sigma}_k(\lambda_x I + \bar{\Sigma}_k)^{-1}(\tilde{x}_k^s - \bar{x}_k) \tag{8}$$

$$\Sigma_k = \bar{\Sigma}_k - \bar{\Sigma}_k(\lambda_x I + \bar{\Sigma}_k)^{-1}\bar{\Sigma}_k, \tag{9}$$

following the Kalman filter equations. Then, it computes in preparation of the next step the predictions $\bar{x}_{k+1} = A\hat{x}_k$ and $\bar{\Sigma}_{k+1} = A\Sigma_k A^T + W$. Note in particular that the estimator implements the equations (5) and (7) between selected samples, which therefore do not need in fact to be reimplemented at the event detector. Another application of the composition theorem leads to the following result.

Theorem 5. At any time $k \geq 0$, the output $\gamma_{0:k}$, $\tilde{x}_{0:k}^s$, $\hat{x}_{0:k}$, $\Sigma_{0:k}$ of the event-triggered sampler is ϵ-differentially private for $\epsilon = \rho n_s(\lambda_\tau + 2\lambda_\nu + \lambda_x)$, where n_s is the number of released samples up to time k (number of times where $\gamma_t = \top$ for $0 \leq t \leq k$).

4. COMPLETING THE ESTIMATOR DESIGN

The task of the estimator is to generate a sequence of state estimates $\hat{x}_k$ given $\tilde{x}^s_{0:k}$. The indication that no sampling was triggered provides some useful information about the state distribution (sometimes called "negative information" in the event-triggered sampling literature), and we describe in this section how we leverage this information in our estimation procedure.

Before the first detected event, given $\bar{x}_0$, $\bar{\Sigma}_0$, and a sequence of observed decisions of no event detected $\gamma_0 = \perp, \ldots, \gamma_k = \perp$, for some $k \geq 0$, the goal of the estimator is to construct at period k a current state estimate $\hat{x}_k$. In fact, we already described with (5) and (7) the equations for our recursive estimation procedure. We complete this description in this section, providing the expression of the parameter η_k, as well as the assumptions and approximations underlying our design.

We start with $k = 0$, and compute first the probability of not sampling, given a value of z_0,

$$\begin{aligned}\mathbb{P}(\gamma_0(z_0) = \perp) &= \mathbb{P}(\nu_0 \geq f_0(z_0) - \tau) \\ &= \int_{\alpha=0}^{\infty} g_0(\alpha, z) p_\tau(\alpha) d\alpha.\end{aligned}$$

Since ν_0 follows a centered Laplace distribution, we have

$$\mathbb{P}(\nu_0 \geq x) = \begin{cases} \frac{1}{2}\exp(-\lambda_\nu x), & x \geq 0, \\ 1 - \frac{1}{2}\exp(\lambda_\nu x), & x < 0. \end{cases}$$

Hence

$$\begin{aligned}\mathbb{P}(\gamma_0(z_0) = \perp) &= \int_{\alpha=0}^{f_0(z_0)} \frac{1}{2} e^{-\lambda_\nu (f_0(z_0)-\alpha)} \lambda_\tau e^{-\lambda_\tau \alpha} d\alpha \\ &+ \int_{\alpha=f_0(z_0)}^{\infty} \left(1 - \frac{1}{2} e^{\lambda_\nu (f_0(z_0)-\alpha)}\right) \lambda_\tau e^{-\lambda_\tau \alpha} d\alpha.\end{aligned}$$

After some calculations, focusing here on the case $\lambda_\nu \neq \lambda_\tau$, we get the likelihood function (a function of z_0)

$$\begin{aligned}&\mathbb{P}(\gamma_0(z_0) = \perp) \\ &= K_\nu(\lambda_\tau, \lambda_\nu) e^{-\lambda_\nu f_0(z_0)} + K_\tau(\lambda_\tau, \lambda_\nu) e^{-\lambda_\tau f_0(z_0)}, \quad (10)\end{aligned}$$

where

$$K_\nu(\lambda_\tau, \lambda_\nu) = \frac{\lambda_\tau}{2(\lambda_\tau - \lambda_\nu)}, \; K_\tau(\lambda_\tau, \lambda_\nu) = \frac{\lambda_\nu^2}{\lambda_\nu^2 - \lambda_\tau^2}.$$

To design our estimator, we now adopt a Bayesian perspective, with (10) providing the expression for $\mathbb{P}(\gamma_0 = \perp | z_0)$. Given a prior distribution $p_0(z)$ for z_0, we can then compute the no-sampling probability

$$\mathbb{P}(\gamma_0 = \perp) = \int_{z \in \mathbb{R}^n} \mathbb{P}(\gamma_0 = \perp | z_0 = z) p_0(z) dz.$$

From the given model, the distribution p_0 should have zero mean and covariance $\bar{\Sigma}_0$. For computational reasons, it is also convenient to assume that it takes a form similar to the likelihood function. This leads us to assume here that $z_0 = \bar{\Sigma}_0^{1/2} y_0$, with $\bar{\Sigma}_0^{1/2}$ the principal square root of $\bar{\Sigma}_0$, and y_0 a vector having independent identically distributed components, with centered Laplace distribution with variance 1. In this case, the vector y_0 has identity covariance matrix and density $p_{y_0}(y) = \frac{1}{2^{n/2}} \exp(-\sqrt{2}\|y\|_1)$, and z_0 is indeed a zero-mean random vector with covariance Σ_0.

By the standard change of variable formula, with Σ_0 invertible, the vector z_0 admits a distribution with density

$$\begin{aligned}p_{z_0}(z) &= \frac{1}{2^{n/2} \left|\det\left(\bar{\Sigma}_0^{1/2}\right)\right|} \exp\left(-\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1\right) \\ &= \frac{1}{\sqrt{\left|\det\left(2\bar{\Sigma}_0\right)\right|}} \exp\left(-\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1\right). \quad (11)\end{aligned}$$

Note that this prior is of the form

$$p_{z_0}(z) = \frac{1}{\sqrt{\left|\det\left(2\bar{\Sigma}_0\right)\right|}} e^{-\sqrt{2}\|\bar{\Sigma}_0^{-1/2}\|_1} f_0(z).$$

Introducing the notation $\lambda_0 := \sqrt{2}\|\bar{\Sigma}_0^{-1/2}\|_1$, the idle probability for this prior is then

$$\begin{aligned}\mathbb{P}(\gamma_0 = \perp) = &\frac{1}{\sqrt{\left|\det\left(2\bar{\Sigma}_0\right)\right|}} \int_{z \in \mathbb{R}^n} K_\nu e^{-(\lambda_\nu/\lambda_0 + 1)\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1} \\ &+ K_\nu e^{-(\lambda_\tau/\lambda_0 + 1)\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1} dz.\end{aligned}$$

To carry out this computation explicitly, note that, for any $\lambda > 0$ and matrix Σ

$$\begin{aligned}&\frac{1}{\sqrt{|\det(2\Sigma)|}} \int_{z \in \mathbb{R}^n} e^{-\lambda\sqrt{2}\|\Sigma^{-1/2} z\|_1} dz \\ &= \frac{1}{\lambda^n \sqrt{|\det(2\Sigma/\lambda^2)|}} \int_{z \in \mathbb{R}^n} e^{-\sqrt{2}\|(\Sigma/\lambda^2)^{-1/2} z\|_1} dz = \frac{1}{\lambda^n},\end{aligned}$$

since the integral of the density (11) is equal to 1. As a result, we get, under the assumed prior for z_0, that

$$\mathbb{P}(\gamma_0 = \perp) = \frac{K_\nu(\lambda_\tau, \lambda_\nu)}{\left(\frac{\lambda_\nu}{\lambda_0} + 1\right)^n} + \frac{K_\tau(\lambda_\tau, \lambda_\nu)}{\left(\frac{\lambda_\tau}{\lambda_0} + 1\right)^n}. \quad (12)$$

The posterior density of z_0 given $\gamma_0 = \perp$ is

$$\begin{aligned}&\mathbb{P}(z_0 = z | \gamma_0 = \perp) = \frac{\mathbb{P}(\gamma_0 = \perp | z_0 = z)\mathbb{P}(z_0 = z)}{\mathbb{P}(\gamma_0 = \perp)} \\ &= \frac{1}{\mathbb{P}(\gamma_0 = \perp)\sqrt{|\det(2\bar{\Sigma}_0)|}} \Big(K_\nu e^{-(\lambda_\nu/\lambda_0 + 1)\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1} \\ &+ K_\nu e^{-(\lambda_\tau/\lambda_0 + 1)\sqrt{2}\|\bar{\Sigma}_0^{-1/2} z\|_1}\Big). \quad (13)\end{aligned}$$

Unfortunately, this posterior density is not of the same form as the prior, as it now involves two terms. We thus follow a suboptimal estimation approach based on certain approximations. First we compute the posterior mean and variance, which can be used in a Kalman filter update providing a *linear* minimum mean-square error (LMMSE) estimate. It is immediate from (13) and the fact that a density of the form (11) has zero mean that

$$\mathbb{E}[z_0 | \gamma_0 = \perp] = 0, \text{i.e., } \mathbb{E}[x_0 | \gamma_0 = \perp] = \bar{x}_0.$$

Hence, the MMSE estimate of x_0 after observing $\gamma_0 = 0$ remains $\bar{x}_0$. Let us now compute the posterior covariance $\mathbb{E}[z_0 z_0^T | \gamma_0 = \perp]$. Again, we use the fact that the density of the form p_0 in (11) has variance $\bar{\Sigma}_0$ to compute, for any $\lambda > 0$ and matrix Σ

$$\begin{aligned}&\frac{1}{\sqrt{|\det(2\Sigma)|}} \int_{z \in \mathbb{R}^n} z z^T e^{-\lambda\sqrt{2}\|\Sigma^{-1/2} z\|_1} dz \\ &= \frac{1}{\lambda^n \sqrt{|\det(2(\Sigma/\lambda^2))|}} \int_{z \in \mathbb{R}^n} z z^T e^{-\sqrt{2}\|(\Sigma/\lambda^2)^{-1/2} z\|_1} dz \\ &= \frac{1}{\lambda^{n+2}} \Sigma.\end{aligned}$$

This gives immediately

$$\mathbb{E}[z_0 z_0^T | \gamma_0 = \perp] = \eta_0 \bar{\Sigma}_0$$

with

$$\eta_0 = \frac{1}{\mathbb{P}(\gamma_0 = 0)} \left(\frac{K_\nu(\lambda_\tau, \lambda_\nu)}{\left(\frac{\lambda_\nu}{\lambda_0} + 1\right)^{n+2}} + \frac{K_\tau(\lambda_\tau, \lambda_\nu)}{\left(\frac{\lambda_\tau}{\lambda_0} + 1\right)^{n+2}} \right). \quad (14)$$

Note that in view of the expression (12) for $\mathbb{P}(\gamma = 0)$, we have $\eta_0 < 1$, as we would expect, i.e., the error covariance decreases after observing $\gamma_0 = \perp$.

For the next period $k = 1$, we have the relation $z_1 = Az_0 + w_0$, which implies, when no sampling occurs ($\gamma_0 = \perp$),

$$\bar{x}_1 := \mathbb{E}[x_1 | \gamma_0 = \perp] = A\bar{x}_0,$$
$$\bar{\Sigma}_1 := \mathbb{E}[z_1 z_1^T | \gamma_0 = \perp] = \eta_0 A \bar{\Sigma}_0 A^T + W. \quad (15)$$

At this point, we can repeat the calculations above for period 1, if we make the following approximation for the distribution of z_1. The conditional idle probability $\mathbb{P}(\gamma_1 = \perp | z_1 = z)$ is given by the expression (10) with f_0 replaced by f_1. The unconditional idle probability is given by (12). Then, to compute the unconditional idle probability and posterior distribution of z_1 given $\gamma_{0:1}$, we approximate the distribution of z_1 given γ_0 by one of the form as (11), with $\bar{\Sigma}_0$ replaced by $\bar{\Sigma}_1$ computed from (15). Under this approximation, the idle probability $\mathbb{P}(\gamma_1 = \perp | \gamma_0 = \perp)$ is given by (12) with λ_0 replaced by $\lambda_1 = \sqrt{2}\|\bar{\Sigma}_1^{-1/2}\|_1$. We also get $\mathbb{E}[z_1 | \gamma_0 = \perp, \gamma_1 = \perp] = 0$, and

$$\mathbb{E}[z_1 z_1^T | \gamma_0 = \perp, \gamma_1 = \perp] = \eta_1 \Sigma_1,$$

with η_1 computed via (14), but with $\mathbb{P}(\gamma_1 = \perp | \gamma_0 = \perp)$ replacing $\mathbb{P}(\gamma_0 = \perp)$ and λ_1 replacing λ_0.

In summary, the approximation involved in the proposed estimator consists in approximating the distribution of z_k given $\gamma_{0:k-1}$ at each "propagation step" of the filter by a distribution of the form (11), with the covariance $\bar{\Sigma}_k$ computed recursively. Note that all the covariance matrix computations are independent of the values of the sensitive data x_k, hence no additional loss of privacy is involved. Figure 2 briefly illustrates the behavior of the sampler and estimator, for the model (1) with $A = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}$, $W = \begin{bmatrix} 0.05 & 0.02 \\ 0.02 & 0.1 \end{bmatrix}$ (the first component of the state trajectories is shown here). An full version of this paper will study the performance of this sampling scheme in details.

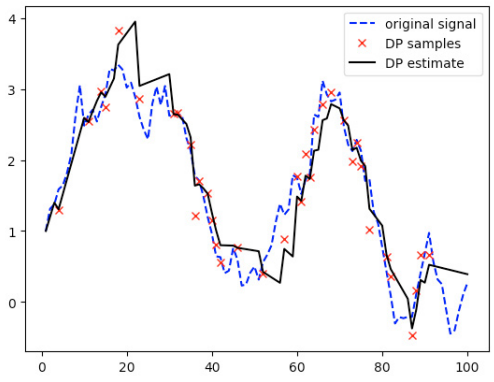

Fig. 2. Example of differentially private signal reconstruction provided by the estimator of the sampling mechanism, with 36 samples selected out of 100, for $\rho = 1$, $\lambda_\nu = 0.2$, $\lambda_\tau = 0.1$, $\lambda_x = 5$.

REFERENCES

Abowd, J.M. (2018). The US Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2867–2867.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third Theory of Cryptography Conference*, 265–284. New York, NY.

Dwork, C., Naor, M., Pitassi, T., and Rothblum, G.N. (2010). Differential privacy under continual observations. In *Proceedings of the ACM Symposium on the Theory of Computing (STOC)*. Cambridge, MA.

Dwork, C., Naor, M., Reingold, O., Rothblum, G.N., and Vadhan, S.P. (2009). On the complexity of differentially private data release: Efficient algorithms and hardness results. In *Symposium on the Theory of Computing*.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4), 211–407.

Erlingsson, Ú., Pihur, V., and Korolova, A. (2014). RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, 1054–1067.

Fan, L. and Xiong, L. (2014). An adaptive approach to real-time aggregate monitoring with differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, 26(9), 2094–2106.

Han, D., Mo, Y., Wu, J., Weerakkody, S., Sinopoli, B., and Shi, L. (2015). Stochastic event-triggered sensor schedule for remote state estimation. *IEEE Transactions on Automatic Control*, 60(10), 2661–2675.

Heemels, W.P.M.H., Johansson, K.H., and Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *Proceedings of the 51st IEEE Conference on Decision and Control (CDC)*.

Le Ny, J. and Mohammady, M. (2018). Differentially private MIMO filtering for event streams. *IEEE Transactions on Automatic Control*, 63(1), 145–157.

Le Ny, J. and Pappas, G.J. (2014). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2), 341–354.

Li, N., Qardaji, W., and Su, D. (2012). On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*.

Lyu, M., Su, D., and Li, N. (2017). Understanding the sparse vector technique for differential privacy. In *Proc. VLDB Endow.*

President's Council of Advisors on Science and Technology (2014). Big data and privacy: A technological perspective. Technical report, Executive Office of the President of the United States.

Wu, J., Ren, X., Hana, D., Shi, D., and Shi, L. (2016). Finite-horizon Gaussianity-preserving event-based sensor scheduling in Kalman filter applications. *Automatica*, 72, 100–107.