Privileged Estimate Fusion With Correlated Gaussian Keystreams

Marko Ristic[1] and Benjamin Noack[1]

Abstract— **Providing cryptographic privacy guarantees in a distributed state estimation problem has been a growing topic of research since the ubiquity of modern public networks. One such guarantee is having different levels of estimation performance achievable by trusted and untrusted users within a sensor network. In the presence of multiple sensor measurements, guaranteeing better estimation performance by the usual means of adding removable noise to measurements is complicated by an alternative for untrusted users to improve their performance: fusing more measurements. Our novel method adds correlated noise at different sensors, restricting the performance gained from fusing additional measurements while guaranteeing better performance to those that can remove it. We extend a cryptographic framework for defining estimation privilege and use this to prove the scheme's security goals, while simulations demonstrate the effects of parameters in a concrete estimation problem. A scheme that can ensure such differences in estimation performance between types of estimators can find applications in priority-based or subscription-based performances in environments where more than one sensor is present.**

I. INTRODUCTION

Sensor data processing and state estimation have long been active areas of research and continue to find applications in modern systems [1], [2]. In the context of distributed sensing environments such as decentralised autonomous vehicles or distributed weather stations, estimation methods relying on Kalman filters and derivatives [3] are particular prevalent due to their recursive, often optimal, estimation properties and their suitability to modelling measurement cross-correlations typically required for data fusion [4], [5]. In recent years, the ubiquity of distributed public networks has seen the additional requirements of preserving algorithm participants' privacies, such as individual contributions or identifying information, become increasingly relevant and has led to an active field of research [6], [7].

While hiding only transmitted information from eavesdroppers can be achieved using common private or public key encryption schemes [8], distributed estimation tasks that preserve participants' privacies require information to remain hidden during partial or complete processing of the task and often justify some leakage [9], [10]. The security goals of these problems are context-specific and have produced a variety of solutions. In [11], non-Bayesian localisation is performed using homomorphic encryption such that individual sensor information and measurements remain private, while in [12], similar goals are achieved in a Bayesian

setting by fusing linear measurements when sensors form a hierarchical network. A combination of homomorphic and order-revealing encryption schemes are used in [9] to solve conservative Gaussian estimate fusion while leaking only the ratios of estimate covariances, and in [10], [13], cryptographic aggregation schemes are introduced and used to leak only total consumptions in an energy grid while hiding individual participant power-usages. In addition to these examples of privacy, optimal estimation performance can itself also be considered leakage, introducing an idea of privilege in estimation, where leakage captures the difference in performance between trusted and untrusted estimators. In the original Global Positioning System (GPS) [14], this was achieved with a secondary encrypted channel that allowed better estimation performance to parties that held an encryption key. Similarly, in [15], a synchronising chaotic system is used to add noise to measurements which can only be removed by estimators knowing its properties. This idea of privileged estimation is further explored in [16], where a formal cryptographic definition of covariance privilege is given and a scheme for a single sensor presented. Here, a synchronised pseudorandom Gaussian keystream adds measurement noise only removable by estimators holding the stream key.

In this work, we consider the definition of estimation privilege presented in [16] and introduce a modified scheme suitable for an environment of multiple sensors, where both holding the secret key and fusing additional measurements can lead to better estimation performance. Our contribution consists of a generalised notion for covariance privilege, the introduction of security requirements for privileged estimation in a well defined multisensor environment and a scheme that satisfies them. Along with a cryptographic proof sketch, simulation results are provided to demonstrate the effects of scheme parameters and how they can be chosen to provide varying amounts of privilege. Use-cases for varying performance in this way include subscription or priority models where some users are provided better results than others. For example, subscription-based weather forecasts using measurements from spatially distributed stations or modular mass-produced sensors that differ in accuracy dependent on cost.

In section II, the multisensor estimation privilege problem is presented. Relevant preliminaries are introduced in section III and the estimation privilege fusion scheme itself in section IV. A cryptographic analysis and the simulation results are then given in sections V and VI, respectively, before the concluding remarks in section VII.

[1]Marko Ristic and Benjamin Noack are with the Autonomous Multisensor Systems Group (AMS), Institute for Intelligent Cooperating Systems (ICS), Otto von Guericke University (OVGU), Magdeburg, Germany `{marko.ristic, benjamin.noack}@ovgu.de`

A. Notation

Lowercase underlined characters $\underline{v}$ are vectors and upper-case bold characters $\mathbf{M}$ are matrices, while $\underline{0}$, $\mathbf{0}$ and $\mathbf{I}$ are the zero vector, zero matrix and identity matrix, respectively, with sizes inferrable from context. $\mathbf{M} \succeq 0$ and $\mathbf{M} \preceq 0$ denote positive and negative semi-definitness, respectively, and $\mathbf{M} \succeq \mathbf{N}$ is short for $\mathbf{M} - \mathbf{N} \succeq 0$. Given an indexed vector $\underline{v}_i$, notation $\underline{v}^{(a:b)}$, $a < b$, will denote the stacked vector $[\underline{v}_a \ \underline{v}_{a+1} \ \cdots \ \underline{v}_b]^\top$, while for indexed matrices $\mathbf{M}_i$, $\mathbf{M}^{(a:b)}$ will denote the block-diagonal matrix with the indexed matrices along the diagonal. Function $\mathsf{Cov}[\cdot]$ computes the covariance of a random vector, $\sim$ denotes distribution, $\dot\sim$ denotes pseudorandom distribution and, in a cryptographic context, $\mathcal{A}(\underline{v})$ denotes the output of an arbitrary algorithm $\mathcal{A}$ given inputs $\underline{v}$.

II. Problem Formulation

We are interested in environments where multiple sensors are present and the fusion of their measurements can lead to better state estimation accuracy of the system they are measuring. In these environments, we want to provide levels of privilege to state estimators such those with a higher privilege can perform better than those with a lower one, while taking into consideration the estimation benefits from fusing additional measurements. We will consider linear and Gaussian models, where a state $\underline{x}_k \in \mathbb{R}^n$, at an integer timestep k, follows a system model given by

$$\underline{x}_k = \mathbf{F}_k \underline{x}_{k-1} + \underline{w}_k, \tag{1}$$

with white noise term $\underline{w}_k \sim \mathcal{N}(\underline{0}, \mathbf{Q}_k)$ and a known covariance $\mathbf{Q}_k \in \mathbb{R}^{n \times n}$. Similarly, measurements $\underline{y}_{k,i} \in \mathbb{R}^m$ from each sensor i, $1 \le i \le N$, follow the measurement models

$$\underline{y}_{k,i} = \mathbf{H}_{k,i} \underline{x}_k + \underline{v}_{k,i}, \tag{2}$$

with white noise term $\underline{v}_{k,i} \sim \mathcal{N}(\underline{0}, \mathbf{R}_{k,i})$ and a known covariance $\mathbf{R}_{k,i} \in \mathbb{R}^{m \times m}$. In addition to these models, we assume that all sensors i are synchronised in timesteps k to simplify later cryptographic evaluation. In practice, this would restrict the presented schemes to scenarios where synchronisation is easier to guarantee, such as ones where measurements are taken infrequently.

Each of the sensors also holds a secret key sk_i, which can be made available to estimators of appropriate privilege. The privileges that we consider, in terms of access to keys and measurements, will be defined by sequential sensors. That is, in the presence of N sensors, we consider exactly N possible privilege levels, where each privilege $p > 0$ corresponds to holding the sequential secret keys sk_j, $1 \le j \le p$, while being unprivileged, $p = 0$, corresponds to holding none. We assume that estimators have access to all "privileged measurements", those from sensors whose keys they hold, and can additionally fuse "unprivileged measurements" from those whose keys they do not hold. To simplify notation, we will consider access to unprivileged measurements to be sequential as well, and can therefore capture their capabilities

by letting $\mathsf{e}^{[p,q]}$ denote an estimator with privilege p and access to measurements from $q \ge p$ sensors i, $1 \le i \le q$.

To cryptographically guarantee the difference between estimation performances, we will use the notion of covariance privilege [16]. As we desire better performance for privileged estimators than unprivileged ones as well as to limit the gained performance from fusing unprivileged measurements, two differences will be considered.

Performance Loss Lower Bound We want to guarantee a lower bound on the estimation performance loss of any unprivileged estimator $\mathsf{e}^{[0,N]}$ on all privilege-p estimators $\mathsf{e}^{[p,p]}$. Naturally, this will remain a lower bound when unprivileged estimators have access to fewer unprivileged measurements or privileged estimators have access to more.

Performance Gain Upper Bound We want to guarantee an upper bound on the estimation performance gain of any estimator $\mathsf{e}^{[p,N]}$ on all privilege-p estimators $\mathsf{e}^{[p,p]}$. Here, the bound remains an upper bound when fewer unprivileged measurements are fused.

The resulting scheme should be such that at least two parameters are responsible for choosing the values of these two bounds.

Remark 1: We stress that the two bounds that will be guaranteed only bound the performances of estimators of the specified forms. That is, nothing is said about estimators which may corrupt sensors to obtain keys beyond their privilege or additional unprivileged measurements. Bounds on leakage caused by corrupting sensors can in some cases be captured by estimators of a new form $\mathsf{e}^{[p',q']}$, but are in general beyond the scope of this work.

III. Preliminaries

When defining our scheme and discussing its cryptographic guarantees, we will make use of Gaussian keystreams and the cryptographic notion of covariance privilege. These have been summarised below.

A. Gaussian keystreams

A Gaussian keystream is a sequence of pseudorandom multivariate Gaussian samples $\underline{g}_k \in \mathbb{R}^m \dot\sim \mathcal{N}(\underline{0}, \mathbf{S})$, $k > 0$, generated using a secret key, for some covariance matrix $\mathbf{S}$. The sequence is indistinguishable from a truly random multivariate Gaussian sequence to any observer who does not hold the key while being reproducible exactly by those that do. The keystream can be constructed from a typical cryptographic stream cipher [8, Ch. 3.6], by first using any common method for random floating-point generation [17] to create a stream of pseudorandom uniform samples $u_x \dot\sim \mathcal{U}(0,1)$, $x > 0$. Here, we make the assumption that floating-point representations of real numbers are sufficiently similar to actual real numbers, made reasonable in practice due to the insignificance of their difference in many applications and their prevalence in estimation theory.

To construct the Gaussian keystream $\underline{g}_k$, $k > 0$, from the uniform one u_x, $x > 0$, a vector of uncorrelated standard

Gaussian samples $\underline{z}_k$ is first generated using the Box-Muller transform,

$$\underline{z}_k = \begin{bmatrix} z_{(k-1)m+1} & \cdots & z_{km} \end{bmatrix}^\top \qquad (3)$$

where

$$z_x = \begin{cases} \sqrt{-2\ln(u_{2x})}\cos(2\pi u_{2x+1}), & x \text{ is odd} \\ \sqrt{-2\ln(u_{2x})}\sin(2\pi u_{2x+1}), & x \text{ is even} \end{cases},$$

and then correlated by

$$\underline{g}_k = \mathbf{S}^{\frac{1}{2}}\underline{z}_k\,, \qquad (4)$$

where $\mathbf{S}^{\frac{1}{2}}$ is any matrix such that $\mathbf{S}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}\top} = \mathbf{S}$.

B. Covariance Privilege

The formal definition of a privileged estimation scheme and accompanying notion of covariance privilege are introduced in [16] and capture a reduction in estimation performance that can always be achieved from a probabilistic polynomial-time (PPT) estimator knowing only unprivileged measurements, and holding no scheme key, to one knowing both the key and privileged measurements. We will use a slight generalisation of this definition to capture an arbitrary difference that can always be achieved between the estimators (rather than only a reduction). A privileged estimation scheme is defined by a pair of probabilistic algorithms (Setup, Noise), given by

Setup$(\mathcal{M}_S, \mathcal{M}_M, \kappa)$ Given system and measurement models $\mathcal{M}_S$ and $\mathcal{M}_M$, and the security parameter κ, public parameters pub and a secrect key sk are returned.

Noise$($pub, sk, $k, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1, \ldots, \underline{y}_k)$ Given public parameters pub, secret key sk, timestep k, models $\mathcal{M}_S$ and $\mathcal{M}_M$, and measurements $\underline{y}_1, \ldots, \underline{y}_k$, modified privileged and unprivileged measurements, $\underline{y}_k^{\{\mathsf{p}\}}$ and $\underline{y}_k^{\{\mathsf{up}\}}$, respectively, are returned.

Using the defintions from [16] for an *estimator* and *negligible covariance* $\mathsf{neglCov}_m(\kappa)$, the notion of covariance privilege is defined as follows.

Definition 3.1: A privileged estimation scheme meets notion $\{\mathbf{D}_1, \mathbf{D}_2, \ldots\}$-*Covariance Privilege for Models* $\mathcal{M}_S$ *and* $\mathcal{M}_M$ if for any PPT estimator $\mathcal{A}$, there exists a PPT estimator $\mathcal{A}'$, such that

$$\begin{aligned} &\mathsf{Cov}\left[\mathcal{A}\left(k, \kappa, \mathsf{pub}, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1^{\{\mathsf{up}\}}, \ldots, \underline{y}_k^{\{\mathsf{up}\}}\right) - \underline{x}_k\right] \\ &- \mathsf{Cov}\left[\mathcal{A}'\left(k, \kappa, \mathsf{pub}, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1^{\{\mathsf{p}\}}, \ldots, \underline{y}_k^{\{\mathsf{p}\}}\right) - \underline{x}_k\right] \\ &\succeq \mathbf{D}_k - \mathsf{neglCov}_m(\kappa) \end{aligned}$$
$$(5)$$

for all $k > 0$, some negligible covariance and where matrices $\mathbf{D}_k$ are semi-definite, *i.e.*, $\mathbf{D}_k \preceq 0$ or $\mathbf{D}_k \succeq 0$. Here, $\mathcal{A}$ and $\mathcal{A}'$ are running in polynomial-time with respect to the parameter κ and all probabilities are taken over randomness introduced in models $\mathcal{M}_S$ and $\mathcal{M}_M$, estimators $\mathcal{A}$ and $\mathcal{A}'$, and algorithms Setup and Noise.

We note that in the generalised notion, definition 3.1, $\mathbf{D}_k \preceq 0$ is allowed, meaning unprivileged estimators may

have lower error covariances (perform better) than privileged ones, albeit by a bounded amount. This feature will be useful when discussing additional estimation performance gainable from fusing unprivileged measurements. Lastly, a sign is also corrected, such that negligible covariances are subtracted (rather than added), correctly capturing the additional negligible performance gainable by an estimator with access to unprivileged measurements.

IV. Privileged Fusion

The idea behind our privileged estimation fusion scheme is to add *correlated* Gaussian keystreams to the measurements from each sensor. These noises can be computed and subtracted by estimators holding respective sensor keys, while their correlation can limit the additional information gained from fusing unprivileged measurements.

A. Noise Generation

Similarly to the Gaussian keystream generation in (4), pseudorandom samples can be correlated in this way even when generated using different stream cipher keys. To parameterise the correlation between noises at each sensor, we introduce a fully correlated component $\mathbf{Z} \in \mathbb{R}^{m \times m}$ and an uncorrelated component $\mathbf{Y} \in \mathbb{R}^{m \times m}$ and define a noise cross-correlation matrix for x noises as $\mathbf{S}^{(x)} \in \mathbb{R}^{xm \times xm}$,

$$\mathbf{S}^{(x)} = \begin{bmatrix} \mathbf{Z} & \cdots & \mathbf{Z} \\ \vdots & \ddots & \vdots \\ \mathbf{Z} & \cdots & \mathbf{Z} \end{bmatrix} + \begin{bmatrix} \mathbf{Y} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y} \end{bmatrix}, \qquad (6)$$

and $\mathbf{S}^{(1)} = \mathbf{Z} + \mathbf{Y}$. The generation of all N multivariate Gaussian noises at timestep k, $\underline{g}_k^{(1:N)}$, can now be written as

$$\underline{g}_k^{(1:N)} = \begin{bmatrix} \underline{g}_{k,1} \\ \vdots \\ \underline{g}_{k,N} \end{bmatrix} = \mathbf{S}^{(N)\frac{1}{2}} \cdot \begin{bmatrix} \underline{z}_{k,1} \\ \vdots \\ \underline{z}_{k,N} \end{bmatrix}, \qquad (7)$$

where each $\underline{z}_{k,i}$ is computed as $\underline{z}_k$ in (3) using uniform samples generated with key sk_i, and $\mathbf{S}^{(N)\frac{1}{2}}$ is a matrix such that $\mathbf{S}^{(N)\frac{1}{2}}\mathbf{S}^{(N)\frac{1}{2}\top} = \mathbf{S}^{(N)}$. Notably, it is important that the vector of the first p noises $\underline{g}_{k,i}$, $1 \leq i \leq p$, in (7), denoted $\underline{g}_k^{(1:p)}$, can be reproduced by an estimator of privilege p, holding only the keys sk_i, $1 \leq i \leq p$. One case where this is possible is when a lower-triangular decomposition, such as the the Cholesky decomposition, is used to compute $\mathbf{S}^{(x)\frac{1}{2}}$ from $\mathbf{S}^{(x)}$. Here, each correlated Gaussian sample $\underline{g}_{k,i}$ is computable from preceeding uniform samples $\underline{z}_{k,j}$, $j \leq i$ only, and the generalised noise generation equation,

$$\underline{g}_k^{(1:p)} = \mathbf{S}^{(p)\frac{1}{2}} \cdot \begin{bmatrix} \underline{z}_{k,1} \\ \vdots \\ \underline{z}_{k,p} \end{bmatrix}, \qquad (8)$$

generates the same first p noises $\underline{g}_k^{(1:p)}$ as would be obtained from (7). This is due to $\mathbf{S}^{(p)\frac{1}{2}} \in \mathbb{R}^{pm \times pm}$ equalling the top left block of matrix $\mathbf{S}^{(N)\frac{1}{2}}$ when using a lower-triangular decomposition.

With (8), at every timestep k, $\underline{g}_k^{(1:N)}$ can be generated using all N keys and used to modify sensor measurements, while the subset $\underline{g}_k^{(1:p)}$ can be generated by estimators of privilege p using only the keys they hold.

B. Measurement Modification

With a way to generate noises for sensors and estimators, we can introduce the means of measurement modification and the observable measurement models for different estimators. Measurement modification is peformed by adding noises $\underline{g}_k^{(1:N)}$ to measurements from each sensor i before making them public, resulting in modified measurement equations for each sensor,

$$\begin{aligned} \underline{y}'_{k,i} &= \underline{y}_{k,i} + \underline{g}_{k,i} \\ &= \mathbf{H}_{k,i}\underline{x}_k + \underline{v}_{k,i} + \underline{g}_{k,i}\,, \end{aligned} \quad (9)$$

with real measurement noise $\underline{v}_{k,i} \sim \mathcal{N}(\underline{0},\mathbf{R}_{k,i})$ and the vector of all added noises $\underline{g}_k^{(1:N)} \overset{\cdot}{\sim} \mathcal{N}(\underline{0},\mathbf{S}^{(N)})$. As we assume that sensors are synchronised, we can capture the correlation between these modified measurements exactly by considering the stacked measurement model for any estimator with access to q measurements, $\mathrm{e}^{[p,q]}$, at time k, given by

$$\begin{aligned} \underline{y}_k'^{(1:q)} &= \underline{y}_k^{(1:q)} + \underline{g}_k^{(1:q)} \\ &= \mathbf{H}_k^{(1:q)}\underline{x}_k + \underline{v}_k^{(1:q)} + \underline{g}_k^{(1:q)} \end{aligned} \quad (10)$$

where $\underline{v}_k^{(1:q)} \sim \mathcal{N}(\underline{0},\mathbf{R}_k^{(1:q)})$ and $\underline{g}_k^{(1:q)} \overset{\cdot}{\sim} \mathcal{N}(\underline{0},\mathbf{S}^{(q)})$, with

$$\underline{y}_k'^{(1:q)} = \begin{bmatrix} \underline{y}'_{k,1} \\ \vdots \\ \underline{y}'_{k,q} \end{bmatrix}, \ \underline{y}_k^{(1:q)} = \begin{bmatrix} \underline{y}_{k,1} \\ \vdots \\ \underline{y}_{k,q} \end{bmatrix}, \ \mathbf{H}_k^{(1:q)} = \begin{bmatrix} \mathbf{H}_{k,1} \\ \vdots \\ \mathbf{H}_{k,q} \end{bmatrix},$$

$$\underline{v}_k^{(1:q)} = \begin{bmatrix} \underline{v}_{k,1} \\ \vdots \\ \underline{v}_{k,q} \end{bmatrix}, \ \mathbf{R}_k^{(1:q)} = \begin{bmatrix} \mathbf{R}_{k,1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_{k,q} \end{bmatrix}$$

and $\mathbf{S}^{(q)} \in \mathbb{R}^{qm\times qm}$ defined by (6).

Since we are using a cryptographically sound stream cipher to generate the added Gaussian keystream, the pseudorandom samples are indistinguishable from truly random ones to estimators without appropriate keys, which leads us to three observable measurement models, *i.e.*, the models that capture all the information available to an estimator exactly, for three types of mutually exhaustive estimators.

Estimators of the form $\mathrm{e}^{[0,q]}$ Here, no keys are held by the unprivileged estimators and the generated noise $\underline{g}_k^{(1:q)}$ is indistinguishable from noise from the truly random distribution $\mathcal{N}(\underline{0},\mathbf{S}^{(q)})$. For these estimators, we can rewrite the measurement equation (10) as the observed measurement model

$$\underline{y}_k^{[0,q]} = \mathbf{H}_k^{(1:q)}\underline{x}_k + \underline{v}_k'\,, \quad (11)$$

with truly Gaussian term $\underline{v}_k' \sim \mathcal{N}(\underline{0},\mathbf{R}_k^{(1:q)}+\mathbf{S}^{(q)})$.

Estimators of the form $\mathrm{e}^{[p,p]}$ Estimators with keys for all the sensors to which they have access can generate

all added noises and subtract them from the recieved measurements. That is, $\underline{g}_k^{(1:p)}$ can be generated and $\underline{y}_k^{[p,p]} = \underline{y}_k'^{(1:p)} - \underline{g}_k^{(1:p)}$ computed to give the observed measurement model equal to recieving unmodified measurements only,

$$\underline{y}_k^{[p,p]} = \mathbf{H}_k^{(1:p)}\underline{x}_k + \underline{v}_k^{(1:p)}\,, \quad (12)$$

where $\underline{v}_k^{(1:p)} \sim \mathcal{N}(\underline{0},\mathbf{R}_k^{(1:p)})$.

Estimators of the form $\mathrm{e}^{[p,q]}$, $p < q$ Lastly, we want the observed measurement model when only some accessible measurements can have have their noises removed. Here, care must be taken when giving the observed model, as the noises from sensors $i > p$, which cannot be removed, are conditionally dependant on the known noises $\underline{g}_k^{(1:p)}$. Since we can generate the noises $\underline{g}_k^{(1:p)}$ and know that $\underline{g}_k^{(1:q)} \overset{\cdot}{\sim} \mathcal{N}(\underline{0},\mathbf{S}^{(q)})$, we can write

$$\begin{aligned} \underline{g}_k^{(1:q)} &= \begin{bmatrix} \underline{g}_k^{(1:p)} \\ \underline{g}_k^{(p+1:q)} \end{bmatrix} \\ &\overset{\cdot}{\sim} \mathcal{N}\left(\begin{bmatrix} \underline{0} \\ \underline{0} \end{bmatrix}, \mathbf{S}^{(q)} = \begin{bmatrix} \mathbf{S}^{(p)} & \mathbf{Z}' \\ \mathbf{Z}'^\top & \mathbf{S}^{(q-p)} \end{bmatrix} \right), \end{aligned} \quad (13)$$

where $\mathbf{Z}' \in \mathbb{R}^{pm\times(q-p)m}$ is a block matrix with every block equal to $\mathbf{Z}$, and compute the conditional pseudorandom Gaussian distribution

$$\begin{aligned} \underline{g}_k^{(p+1:q)} &\mid \underline{g}_k^{(1:p)} \\ &\overset{\cdot}{\sim} \mathcal{N}\Big(\mathbf{Z}'^\top\mathbf{S}^{(p)-1}\underline{g}_k^{(1:p)}, \\ &\quad \mathbf{S}^{(q-p)} - \mathbf{Z}'^\top\mathbf{S}^{(p)-1}\mathbf{Z}'\Big). \end{aligned} \quad (14)$$

Now, subtracting the known noises $\underline{g}_k^{(1:p)}$ and the means of the unknown noises (14) from recieved measurements,

$$\underline{y}_k^{[p,q]} = \underline{y}_k'^{(1:q)} - \begin{bmatrix} \underline{g}_k^{(1:p)} \\ \mathbf{Z}'^\top\mathbf{S}^{(p)-1}\underline{g}_k^{(1:p)} \end{bmatrix}, \quad (15)$$

and accounting for unknown pseudorandom noises being indistinguishable from random, a zero-mean observed measurement model can be written as

$$\underline{y}_k^{[p,q]} = \mathbf{H}_k^{(1:q)}\underline{x}_k + \underline{v}_k' \quad (16)$$

where

$$\underline{v}_k' \sim \mathcal{N}\Bigg(\underline{0}, \\ \begin{bmatrix} \mathbf{R}_k^{(1:p)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{(q-p)} - \mathbf{Z}'^\top\mathbf{S}^{(p)-1}\mathbf{Z}' + \mathbf{R}_k^{(p+1:q)} \end{bmatrix}\Bigg).$$

Remark 2: Recalling that access to unprivileged measurements are sequential to simplify notation, (13), (15) and (16) can be generalised when having access to arbitrary $q-p$ non-sequential unprivileged measurements $\underline{y}_{k,i}$, $p < i \le q$, by appropriately replacing elements $\mathbf{Z}'$ and $\mathbf{S}^{(q-p)}$ in (13).

From the observed measurement models (11), (12) and (16) we can tell that the parameters $\mathbf{Z}$ and $\mathbf{Y}$ (within matrices

$\mathbf{S}^{(x)}$) will control the difference in estimation performance between the three types of estimators. The two differences we wish to cryptographically guarantee from section II, and how $\mathbf{Z}$ and $\mathbf{Y}$ affect them, will be more formally explored in sections V and VI.

C. Noise Distribution

While we have described a method for generating noises that modify N measurements and result in different observed measurement models depending on estimator privilege, we have not discussed where the noise is generated and how it is distributed to sensors. To handle the inherent correlation of the noises $g_k^{(1:N)}$, they can be generated either centrally before distribution to sensors or sequentially at the sensors themselves, given previously generated values.

Central noise generation To compute noises centrally, (8) can be computed for all N noises at a central processor and each noise $\underline{g}_{k,i}$ sent to the respective sensor i before it modifies its local measurement by (9).

Sequential noise generation To compute the same noises sequentially for each timestep k, sensor 1 can generate its noise independently using its current standard Gaussian sample $\underline{z}_{k,1}$, by $\underline{g}_{k,1} = \mathbf{S}^{(1)\frac{1}{2}}\underline{z}_{k,1}$. Each following sensor $i > 1$ can generate its noise $\underline{g}_{k,i}$ given the preceeding noises $g_k^{(1:i-1)}$, following the conditional reasoning in (14), as

$$\underline{g}_{k,i} = \mathbf{Z}'^{\top}\mathbf{S}^{(i-1)-1}\underline{g}_k^{(1:i-1)} + \\ (\mathbf{S}^{(1)} - \mathbf{Z}'^{\top}\mathbf{S}^{(i-1)-1}\mathbf{Z}')^{\frac{1}{2}}\underline{z}_{k,i}. \quad (17)$$

After local noise generation, sensor i sends its and preceding noises, $g_k^{(1:i)}$, to the next sensor $i + 1$. This method has the clear downside of increasing communication costs with each successive generation but requires no central communicator.

In both cases above, the computation of all noises $\underline{g}_k^{(1:N)}$ can be performed offline, reducing the complexity of real-time measurement modification.

V. CRYPTOGRAPHIC BOUNDS

To give proof sketches of the cryptographic guarantees provided by the presented scheme, we first recall some assumptions. We consider floating-point numbers to be sufficiently close to real random numbers that real-number proofs still hold, and that all sensors are synchronised in k such that observed measurement models (11), (12) and (16) are exactly correct. Using these assumptions, the proofs rely on the optimality of the linear Kalman filter (KF) [3] to produce a series of covariances for optimal estimators before taking their differences to obtain the *Performance Loss Lower Bound* and *Performance Gain Upper Bound* from section II. Similarly to [16], these series can be used as $\mathbf{D}_1, \mathbf{D}_2, \ldots$ in definition 3.1 for appropriately formulated privileged estimation schemes for the two bounds, and demonstrate that the existence of an estimator violating the notion implies the existence of a linear estimator with error covariance lower than the KF. This guarantees the bound by contrapositive.

A. Performance Loss Lower Bound (PLLB)

First, we consider the lower bound to the loss in estimation performance an estimator $e^{[0,N]}$ has over all estimators $e^{[p,p]}$. Since the observed measurement models for these estimators, (11) and (12), interpret available measurements as a single stacked measurement, and since we do not consider estimators that corrupt sensors, we can treat the stacked measurement as coming from a single sensor and use the notion of covariance privilege in definition 3.1 to guarantee the bound. The associated privileged estimation scheme for the PLLB can be written for each privilege p as

Setup Given the system model (1), all measurements models (2) (interpretable as a single stacked measurement model) and a security parameter κ used by all sensors, generate N stream cipher keys sk_i, $1 \leq i \leq N$, and let the secret key sk include all N keys. Generate the correlated and uncorrelated noise components $\mathbf{Z}$ and $\mathbf{Y}$, an initial estimate and error covariance $\hat{\underline{x}}_0$ and $\mathbf{P}_0$, and include these in the public parameters pub.

Noise$_{\mathsf{PLLB}}$ Given parameters, cipher keys, a timestep k and true sensor measurements $y_k^{(1:N)}$, let $\underline{y}_k^{\{\mathsf{up}\}} = \underline{y}_k^{[0,N]}$ following (11) and $\underline{y}_k^{\{\mathsf{p}\}} = \underline{y}_k^{[p,p]}$ following (12).

With the above formulation, we can use the KF to recursively compute the optimal estimate error covariances for estimators with access to only measurements $\underline{y}_k^{\{\mathsf{up}\}}$ or $\underline{y}_k^{\{\mathsf{p}\}}$, for all k. Since the KF preserves initial covariance order, $\mathbf{P}_k \succeq \mathbf{P}_k' \implies \mathbf{P}_{k+1} \succeq \mathbf{P}_{k+1}'$, a lower bound can be guaranteed when using the initial covariance $\mathbf{P}_0 = \mathbf{0}$. Therefore, the minimum achievable error covariance for an estimator $e^{[0,N]}$, with access to measurements $\underline{y}_k^{\{\mathsf{up}\}} = \underline{y}_k^{[0,N]}$, is given by the combined KF predict and update equations

$$\mathbf{P}_k^{[0,N]} = \Big(\mathbf{I} - (\mathbf{F}_k\mathbf{P}_{k-1}^{[0,N]}\mathbf{F}_k^{\top} + \mathbf{Q}_k)\mathbf{H}_k^{(1:N)\top} \cdot \\ \big(\mathbf{H}_k^{(1:N)}(\mathbf{F}_k\mathbf{P}_{k-1}^{[0,N]}\mathbf{F}_k^{\top} + \mathbf{Q}_k)\mathbf{H}_k^{(1:N)\top} \\ + \mathbf{R}_k^{(1:N)} + \mathbf{S}^{(N)}\big)^{-1}\mathbf{H}_k^{(1:N)}\Big) \cdot \\ \Big(\mathbf{F}_k\mathbf{P}_{k-1}^{[0,N]}\mathbf{F}_k^{\top} + \mathbf{Q}_k\Big), \quad (18)$$

when $\mathbf{P}_0^{[0,N]} = \mathbf{0}$. Similarly, the same can be done for an estimator $e^{[p,p]}$, with access to measurements $\underline{y}_k^{\{\mathsf{p}\}} = \underline{y}_k^{[p,p]}$, as

$$\mathbf{P}_k^{[p,p]} = \Big(\mathbf{I} - (\mathbf{F}_k\mathbf{P}_{k-1}^{[p,p]}\mathbf{F}_k^{\top} + \mathbf{Q}_k)\mathbf{H}_k^{(1:p)\top} \cdot \\ \big(\mathbf{H}_k^{(1:p)}(\mathbf{F}_k\mathbf{P}_{k-1}^{[p,p]}\mathbf{F}_k^{\top} + \mathbf{Q}_k)\mathbf{H}_k^{(1:p)\top} \\ + \mathbf{R}_k^{(1:p)}\big)^{-1}\mathbf{H}_k^{(1:p)}\Big) \cdot \\ \Big(\mathbf{F}_k\mathbf{P}_{k-1}^{[p,p]}\mathbf{F}_k^{\top} + \mathbf{Q}_k\Big) \quad (19)$$

and $\mathbf{P}_0^{[p,p]} = \mathbf{0}$. The bounds (18) and (19) are constructed such that at every timestep k,

$$\mathbf{P}_k^{[0,N]} \preceq \\ \mathsf{Cov}\Big[\mathcal{A}\Big(k, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1^{[0,N]}, \ldots, \underline{y}_k^{[0,N]}\Big) - \underline{x}_k\Big] \quad (20)$$

and

$$\mathbf{P}_k^{[p,p]} \preceq \\ \mathsf{Cov}\left[\mathcal{A}\left(k, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1^{[p,p]}, \ldots, \underline{y}_k^{[p,p]} \right) - \underline{x}_k \right] \quad (21)$$

hold. Therefore, taking their difference,

$$\mathbf{D}_{\mathsf{PLLB},k} = \mathbf{P}_k^{[0,N]} - \mathbf{P}_k^{[p,p]}, \quad (22)$$

produces a series such that for any PPT estimator $\mathsf{e}^{[0,N]}$, an equivalent PPT estimator $\mathsf{e}^{[p,p]}$, lower bounded in error by (19), can always be created such that the difference between their error covariances at time k is $\mathbf{D}_{\mathsf{PLLB},k}$. Here, the PPT requirement is used to guarantee the indistinguishabilty of the pseudorandom streams to truly random ones. The existence of an estimator of the form $\mathsf{e}^{[0,N]}$ where this condition cannot be met, implies the existence of an estimator with error covariances smaller than the KF for linear models. As no such estimator exists, we conclude that the Setup and $\mathsf{Noise}_{\mathsf{PLLB}}$ algorithms above meet $\{\mathbf{D}_{\mathsf{PLLB},1}, \mathbf{D}_{\mathsf{PLLB},2}, \ldots\}$-*Covariance Privilege for System Model* (1) *and Stacked Measurement Models* (2).

In the above, we lower bound the estimation performance loss an estimator $\mathsf{e}^{[0,N]}$ has over all estimators $\mathsf{e}^{[p,p]}$. In the cases where the unprivileged estimator has access to fewer measurements, $\mathsf{e}^{[0,q]}$, $q < N$, or the privileged one to more, $\mathsf{e}^{[p,q]}$, $q > p$, the achievable difference can only increase (fewer measurements can only increase error covariance while more can only decrease it). This ensures the computed bound remains a lower bound for *any* unprivileged estimator over all privilege-p ones.

B. Performance Gain Upper Bound (PGUB)

Similar to the lower bound above, we can use the same properties of the KF to give an upper bound to the gain in estimation performance an estimator $\mathsf{e}^{[p,N]}$ has on all estimators $\mathsf{e}^{[p,p]}$. The associated privileged estimation scheme for the PGUB for each privilege p is given by the same Setup algorithm as in section V-A and

$\mathsf{Noise}_{\mathsf{PGUB}}$ Given parameters, cipher keys, a timestep k and true sensor measurements $\underline{y}_k^{(1:N)}$, let $\underline{y}_k^{\{\mathsf{up}\}} = \underline{y}_k^{[p,N]}$ following (16) and $\underline{y}_k^{\{\mathsf{p}\}} = \underline{y}_k^{[p,p]}$ following (12).

The minimum error covariances achievable by an estimator $\mathsf{e}^{[p,p]}$ is again given by (19) and $\mathbf{P}_0^{[p,p]} = \mathbf{0}$. For an estimator $\mathsf{e}^{[p,N]}$ with access to measurements $\underline{y}_k^{\{\mathsf{up}\}} = \underline{y}_k^{[p,N]}$ it is given by

$$\mathbf{P}_k^{[p,N]} = \Big(\mathbf{I} - (\mathbf{F}_k \mathbf{P}_{k-1}^{[p,N]} \mathbf{F}_k^\top + \mathbf{Q}_k) \mathbf{H}_k^{(1:N)\top} \cdot \\ \big(\mathbf{H}_k^{(1:N)} (\mathbf{F}_k \mathbf{P}_{k-1}^{[p,N]} \mathbf{F}_k^\top + \mathbf{Q}_k) \mathbf{H}_k^{(1:N)\top} \quad (23) \\ + \mathbf{X} \big)^{-1} \mathbf{H}_k^{(1:N)} \Big) \big(\mathbf{F}_k \mathbf{P}_{k-1}^{[p,N]} \mathbf{F}_k^\top + \mathbf{Q}_k \big),$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{R}_k^{(1:p)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{(N-p)} - \mathbf{Z}'^\top \mathbf{S}^{(p)-1} \mathbf{Z}' + \mathbf{R}_k^{(p+1:N)} \end{bmatrix} \quad (24)$$

and $\mathbf{P}_0^{[p,N]} = \mathbf{0}$. Again, the bounding series are such that (21) and

$$\mathbf{P}_k^{[p,N]} \preceq \\ \mathsf{Cov}\left[\mathcal{A}\left(k, \mathcal{M}_S, \mathcal{M}_M, \underline{y}_1^{[p,N]}, \ldots, \underline{y}_k^{[p,N]} \right) - \underline{x}_k \right] \\ (25)$$

hold. Taking their difference,

$$\mathbf{D}_{\mathsf{PGUB},k} = \mathbf{P}_k^{[p,N]} - \mathbf{P}_k^{[p,p]}, \quad (26)$$

produces a series such that for any PPT estimator $\mathsf{e}^{[p,N]}$, an equivalent PPT estimator $\mathsf{e}^{[p,p]}$, lower bounded in error by (19), can always be created such that the difference between their error covariances at time k is given by $\mathbf{D}_{\mathsf{PGUB},k}$. With the same reasoning as for the lower bound, we conclude that the Setup and $\mathsf{Noise}_{\mathsf{PGUB}}$ algorithms above meet $\{\mathbf{D}_{\mathsf{PGUB},1}, \mathbf{D}_{\mathsf{PGUB},2}, \ldots\}$-*Covariance Privilege for System Model* (1) *and Stacked Measurement Models* (2).

Here, $\mathbf{D}_{\mathsf{PGUB},k} \preceq \mathbf{0}$ for all $k > 0$ and, by the above definition, lower bounds the (negative) loss in performance an estimator $\mathsf{e}^{[p,N]}$ has over all estimators $\mathsf{e}^{[p,p]}$. We refer to the bound as an upper bound as its negation $-\mathbf{D}_{\mathsf{PGUB},k}$, $k > 0$, upper bounds the estimation performance gain achievable by $\mathsf{e}^{[p,N]}$ over all $\mathsf{e}^{[p,p]}$, as desired in section II. In the case where fewer unprivileged measurements are accessible, $\mathsf{e}^{[p,q]}$, $q < N$, this gain can only decrease, keeping the upper bound valid for any estimators $\mathsf{e}^{[p,q]}$, $q > p$.

VI. SIMULATION

In addition to showing how the estimation performance loss and gain bounds can be computed, we have simulated optimal estimators to demonstrate the effects of the correlated and uncorrelated components, $\mathbf{Z}$ and $\mathbf{Y}$, respectively. Simulations were implemented in the Python programming language using a constant velocity system model with location measurements, given by parameters

$$\mathbf{F}_k = \begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (27)$$

$$\mathbf{Q}_k = 10^{-3} \cdot \begin{bmatrix} 0.42 & 1.25 & 0 & 0 \\ 1.25 & 5 & 0 & 0 \\ 0 & 0 & 0.42 & 1.25 \\ 0 & 0 & 1.25 & 5 \end{bmatrix}, \quad (28)$$

$$\mathbf{H}_{k,i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } \mathbf{R}_{k,i} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}, \quad (29)$$

for all k and sensors i, $1 \leq i \leq N = 4$. The considered correlated and uncorrelated parameters were restricted to the forms $\mathbf{Z} = \Sigma_z \cdot \mathbf{I}$ and $\mathbf{Y} = \Sigma_y \cdot \mathbf{I}$ for simplicity and all estimators executed a linear Kalman filter with the parameters above and exact knowledge of the initial state ($\mathbf{P}_0 = \mathbf{0}$).

Figure 1 shows the errors for different privilege estimators when added noise parameters Σ_z and Σ_y are held constant. As would be expected, error decreases when more keys

are available, while a further decrease is achieved as more additional unprivileged measurements are fused. Here, the difference in mean squared error (MSE) between $e^{[0,4]}$ and $e^{[p,p]}$ (blue and orange), and between $e^{[p,p]}$ and $e^{[p,4]}$ (green and orange), are bounded on average by the PLLB and PGUB, respectively, when computed for each privilege p with $\Sigma_z = 2$ and $\Sigma_y = 10$.
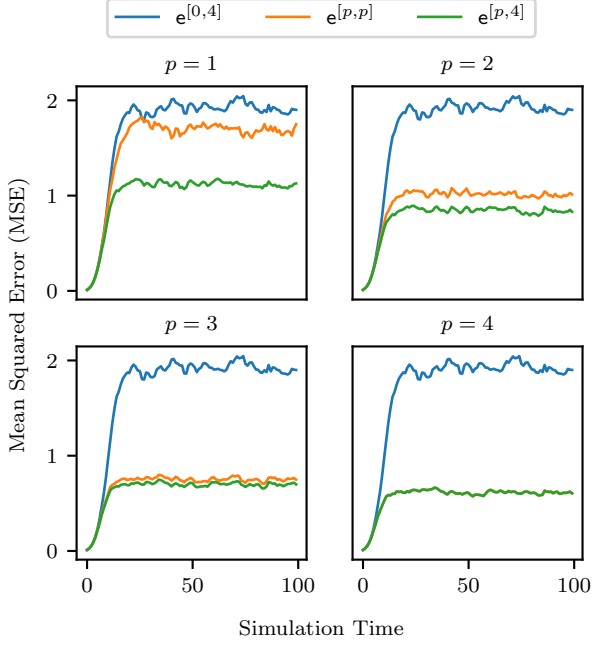

Fig. 1. Average error of 1000 simulation runs for different estimators when $\Sigma_z = 2$ and $\Sigma_y = 10$.

To demonstrate the effect of parameters Σ_z and Σ_y (and therefore $\mathbf{Z}$ and $\mathbf{Y}$), figure 2 shows their effect on the MSE given fixed estimators. It can be seen that Σ_z has a more prominent effect on the PLLB while Σ_y has it on the PGUB. However, it can also be observed that both parameters affect both bounds to some degree, revealing some limitations when specific bounds are desired using the proposed scheme. Figure 3 further captures this relation between the bounds and the parameters Σ_z and Σ_y. As the simulated system is asymptotically stable, steady state error covariances are reached as $k \to \infty$, and therefore $\mathbf{D}_{\mathsf{PLLB},k}$ and $\mathbf{D}_{\mathsf{PGUB},k}$ stabilise as well. From the plot, we can see that increasing the fully correlated noise parameter Σ_z cannot greatly reduce the PGUB (*i.e.* bring $\mathsf{tr}(\mathbf{D}_{\mathsf{PGUB},k})$ closer to 0), likely due to privileged estimators estimating this component of the noise well and the remaining uncorrelated component staying unchanged. Simultaneously, however, the fully correlated component can greatly increase the PLLB (*i.e.* take $\mathsf{tr}(\mathbf{D}_{\mathsf{PLLB},k})$ further from 0) as it increases the redundancy of fusing only unprivileged measurements. The effects of Σ_y are less one-sided, as uncorrelated noise still affects fusing only unprivileged measurements and increases the PLLB, albeit less drastically. Its increase has a stronger effect in reducing the PGUB to near 0, as enough uncorrelated noise will make fusing additional unprivileged measurements hold
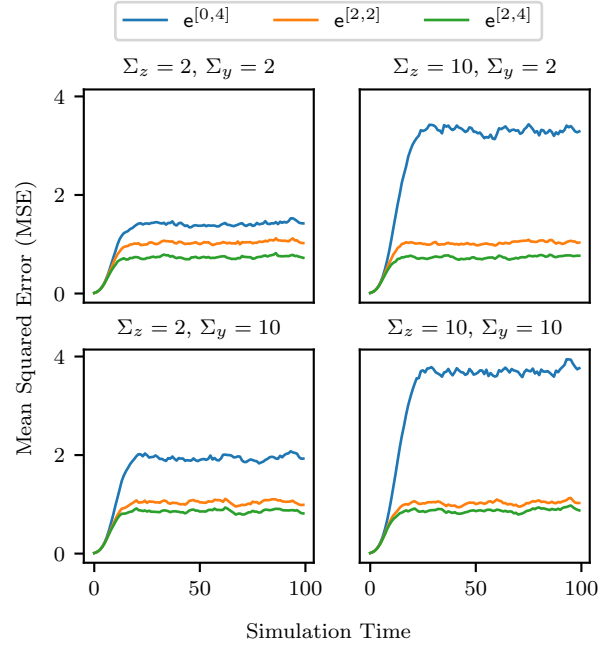

Fig. 2. Average error of 1000 simulation runs of unprivileged and privilege-2 estimators for varying Σ_z and Σ_y.

little additional information even when some keys are known.

Figure 3 also shows how the bounds are affected by the privilege p they are computed for. As can be expected, a higher privilege p results in fewer possible unprivileged measurements to fuse, resulting in a lower PGUB, while a lower p can lead to unprivileged estimators performing better than privileged ones when the added noise is small (negative definite $\mathbf{D}_{\mathsf{PLLB},k}$ when $p = 1$ and $\Sigma_y = 5$).

VII. Conclusion

The presented method demonstrates how different levels of estimation performance can be cryptographically guaranteed in a network of multiple sensors. The problem considered requires sequential access to sensors and sensor keys and allows for two free parameters, $\mathbf{Z}$ and $\mathbf{Y}$, to loosely control the two relevant cryptographic bounds. The different bounds for each privilege, changing bounds over time and the relation between these parameters and the bounds themselves mean that choosing these parameters is a task-specific problem where care must be taken to meet any desired bounds without overly affecting others. Resulting cryptographic bounds can, however, always be computed exactly.

Future work on this topic includes deriving parameters that affect the relevant cryptographic bounds independently, relaxing sequential sensor and key access requirements and exploring methods for decentralised correlated noise generation with fewer communication costs.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Dover Publications, 1979.
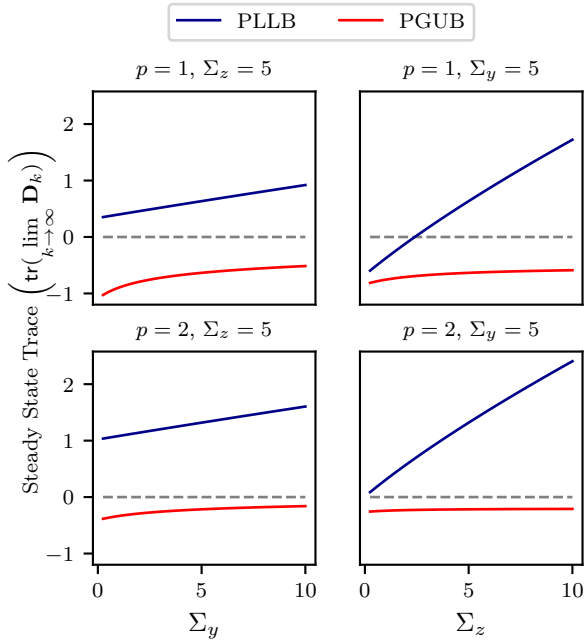[2] D. Simon, *Optimal State Estimation: Kalman, H Infinity and Nonlinear Approaches*. Wiley-Interscience, 2006.

Fig. 3. Steady state traces of the PLLB and PGUB for privileges $p = 1$ and $p = 2$ when Σ_z and Σ_y are varied.

[3] A. J. Haug, *Bayesian Estimation and Tracking: A Practical Guide*. John Wiley & Sons, 2012.

[4] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*. CRC press, 1998.

[5] M. Liggins, C. Y. Chong, D. Hall, and J. Llinas, *Distributed Data Fusion for Network-Centric Operations*. CRC Press, 2012.

[6] K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," *IEEE Internet Computing*, vol. 16, no. 1, pp. 69–73, 2012.

[7] M. Brenner, J. Wiebelitz, G. von Voigt, and M. Smith, "Secret Program Execution in the Cloud Applying Homomorphic Encryption," in *5th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*, 2011, pp. 114–119.

[8] J. Katz and Y. Lindell, *Introduction to Modern Cryptography: Principles and Protocols*. Chapman & Hall, 2008.

[9] M. Ristic, B. Noack, and U. D. Hanebeck, "Secure Fast Covariance Intersection Using Partially Homomorphic and Order Revealing Encryption Schemes," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 217–222, 2021.

[10] E. Shi, T.-H. H. Chan, and E. Rieffel, "Privacy-Preserving Aggregation of Time-Series Data," *Annual Network & Distributed System Security Symposium (NDSS)*, p. 17, 2011.

[11] A. Alanwar, Y. Shoukry, S. Chakraborty, P. Martin, P. Tabuada, and M. Srivastava, "PrOLoc: Resilient Localization with Private Observers Using Partial Homomorphic Encryption," in *16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2017, pp. 41–52.

[12] M. Aristov, B. Noack, U. D. Hanebeck, and J. Müller-Quade, "Encrypted Multisensor Information Filtering," in *21st International Conference on Information Fusion (Fusion 2018)*, 2018, pp. 1631–1637.

[13] M. Joye and B. Libert, "A Scalable Scheme for Privacy-Preserving Aggregation of Time-Series Data," in *International Conference on Financial Cryptography and Data Security*, ser. Lecture Notes in Computer Science. Springer, 2013, pp. 111–125.

[14] P. D. Groves, "Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 2, pp. 26–27, 2015.

[15] C. Murguia, I. Shames, F. Farokhi, and D. Nešić, "Information-Theoretic Privacy Through Chaos Synchronization and Optimal Additive Noise," in *Privacy in Dynamical Systems*. Springer, 2020, pp. 103–129.

[16] M. Ristic, B. Noack, and U. D. Hanebeck, "Cryptographically Privileged State Estimation With Gaussian Keystreams," *IEEE Control Systems Letters*, vol. 6, pp. 602–607, 2022.

[17] F. Goualard, "Generating Random Floating-Point Numbers by Dividing Integers: A Case Study," *International Conference on Computational Science (ICCS)*, vol. 12138, pp. 15–28, 2020.