A Randomized Filtering Strategy Against Inference Attacks on Active Steering Control Systems

Ehsan Nekouei, *Member, IEEE*, Mohammad Pirani, Henrik Sandberg, *Senior Member, IEEE*, and Karl H. Johansson, *Fellow, IEEE*

***Abstract*— In this paper, we develop a framework against inference attacks aimed at inferring the values of the controller gains of an active steering control system (ASCS). We first show that an adversary with access to the shared information by a vehicle, via a vehicular ad hoc network (VANET), can reliably infer the values of the controller gains of an ASCS. This vulnerability may expose the driver as well as the manufacturer of the ASCS to severe financial and safety risks. To protect controller gains of an ASCS against inference attacks, we propose a randomized filtering framework wherein the lateral velocity and yaw rate states of a vehicle are processed by a filter consisting of two components: a nonlinear mapping and a randomizer. The randomizer randomly generates a pair of pseudo gains which are different from the true gains of the ASCS. The nonlinear mapping performs a nonlinear transformation on the lateral velocity and yaw rate states. The nonlinear transformation is in the form of a dynamical system with a feedforward-feedback structure which allows real-time and causal implementation of the proposed privacy filter. The output of the filter is then shared via the VANET. The optimal design of randomizer is studied under a privacy constraint that determines the protection level of controller gains against inference attacks, and is in terms of mutual information. It is shown that the optimal randomizer is the solution of a convex optimization problem. By characterizing the distribution of the output of the filter, it is shown that the statistical distribution of the filter's output depends on the pseudo gains rather than the true gains. Using information-theoretic inequalities, we analyze the inference ability of an adversary in estimating the control gains based on the output of the filter. Our analysis shows that the performance of any estimator in recovering the controller gains of an ASCS based on the output of the filter is limited by the privacy constraint. The performance of the proposed privacy filter is compared with that of an additive noise privacy mechanism. Our numerical results show that the proposed privacy filter significantly outperforms the additive noise mechanism, especially in the low distortion regime.**

Manuscript received February 19, 2021; revised August 12, 2021 and October 25, 2021; accepted November 3, 2021. Date of publication November 23, 2021; date of current version December 16, 2021. The work was supported in part by the Research Grants Council of Hong Kong under the Project CityU 9048210, in part by the Chow Sang Sang Group Research Fund Sponsored by Chow Sang Sang Holdings International Ltd., in part by the Knut and Alice Wallenberg Foundation, in part by the Swedish Foundation for Strategic Research, and in part by the Swedish Research Council. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kai Zeng. *(Corresponding author: Ehsan Nekouei.)*

Ehsan Nekouei is with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong (e-mail: enekouei@cityu.edu.hk).

Mohammad Pirani is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 1A1, Canada (e-mail: mohammadpirani@utoronto.ca).

Henrik Sandberg and Karl H. Johansson are with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden, and also with Digital Futures, SE-100 44 Stockholm, Sweden (e-mail: hsan@kth.se; kallej@kth.se).

Digital Object Identifier 10.1109/TIFS.2021.3130439

***Index Terms*— Information privacy and security, inference attack, vehicular ad hoc networks (VANETs), active steering control system (ASCS), randomized filtering.**

I. Introduction

A. Motivation

VEHICULAR communication systems play a critical role in intelligent transportation systems by enabling vehicles-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Information exchange via vehicular networks enables services such as cooperative collision warning, incident management, platooning, and traffic information notification. The information of individual vehicles, including vehicle's kinematic, dynamic, and geometric parameters, are disseminated through the vehicular network in order to enhance the active safety of vehicles (e.g., collision detection, lane changing warning, and cooperative merging), increasing the traffic throughput, as well as infotainment applications (e.g., interactive gaming, and file and other valuable information sharing). All these tasks can be easily done among vehicles in proximity, or vehicles multiple hops away in a vehicular ad hoc network (VANET) without the assistance of any built infrastructure.

Although information exchange via V2V and V2I systems provides numerous benefits, it exposes the control systems of vehicles at the risk of *privacy breach via inference attacks*. More precisely, one can identify the parameters of control systems, *e.g.,* controller gains of vehicular control systems, based on the shared information. The privacy breach of control parameters exposes the system designers to serious risks. Each vehicle manufacturer dedicates numerous theoretical and experimental efforts to designing and optimizing vehicular control systems. Thus, from the manufacturers' perspective, it is vital to keep the design parameters private in order to protect the right of usage and a privacy breach of sensitive parameters exposes them to severe financial risks.

Privacy breaches of control parameters might have dangerous consequences for the drivers. It has been demonstrated experimentally in [1] and [2] that modern vehicles are highly vulnerable to cyber-physical attacks. An attacker can launch detrimental attacks on vehicles using the knowledge of the parameters of its control systems. An example of such model-based stealthy attacks is the zero dynamics in which the adversary conceals the attack signal in the so-called output-nulling space, even if a large amount of false data are injected into the plant [3]–[6]. With this in mind, any privacy breach of the vehicle's dynamical model, including controller gains, exposes the drivers to highly impactful cyber-physical attacks.

The active steering control system (ASCS), a safety critical control unit of a vehicle, ensures the vehicle's lateral stability during maneuvers such as lane keeping. Under the existing communications standards, *e.g.*, SAE J2735, the states of the ASCS and the driver steering command are shared via vehicular communication networks which may result in the privacy breach of controller gains of a vehicle's ASCS. Consequently, the ASCS can be compromised by a malicious agent using model-based attacks (via CAN Bus) which is severely dangerous especially in obstacle avoidance at high speeds. Motivated by these observations, this paper investigates the privacy aspect of controller gains of ASCSs under vehicular communications.

B. Related Work

The leakage of private information via sensor measurements as well as various solutions for ensuring privacy in such scenarios have been investigated in the literature, *e.g.,* [7]–[10]. Li and Oechtering in [11] considered a multi-sensor hypothesis testing problem and proposed a privacy-aware decision fusion rule that minimizes the Bayes risk subject to a constraint on the inference capability of an adversary with access to the local decisions of a subset of sensors. The authors in [12] studied the optimal privacy-aware design of the Neyman-Pearson test under a set-up similar to that of [11].

Information-theoretic approach to data privacy has been extensively studied in the literature, *e.g.,* [13]–[17] and references therein. In this line of work, the privacy filter is designed such that the distortion between the input and output of the filter is minimized subject to a privacy constrained captured by information theoretic notions such as conditional entropy. The authors in [18] considered a discrete-time Markov chain which carries public information and is correlated with a private Markov chain. They studied the optimal design of privacy filter for the public chain when the filter has access to the outputs of both chains. The author in [19] studied the optimal control of a Markov decision process in presence of an adversary that is curious about the state of the process and has access to the input and output of the process. The interested reader is referred to [20] for an overview of information-theoretic approaches to privacy in estimation and control.

Variations of differential privacy have been used in the literature to develop privacy-aware solutions for estimation, filtering and average consensus problems. In this approach, a randomized mechanism perturbs privacy-sensitive data, typically by adding noise, prior to sharing the data with an honest-but-curious adversary. Ny and Pappas in [21] proposed the notion of differentially private filtering to ensure the privacy of the measurements of a dynamical system. The authors in [22] studied the state estimation problem in a distribution power network under the privacy constraints of the consumers. The authors of [23] and [24] proposed privacy-aware algorithms for the average consensus problem to ensure the privacy of initial states of agents.

Wang *et al.* considered a distributed multi-agent control problem in [25] and developed a differential privacy scheme to ensure the privacy of the initial state and the preferred target way-points of each agent. The privacy filter design problem for the output measurements of a linear Gaussian system was studied in [26]. It was shown that the nonlinear transformation entails two one-step ahead Kalman predictors and requires the knowledge of all the past inputs and outputs of the privacy filter. Different from [26], we study the privacy filter design problem when the privacy filter has access to the states of a linear system (rather than its outputs) when the process noise is arbitrarily distributed. We show that, in our set-up, the nonlinear transformation is in the form of a dynamical system where the output of the filter at each time-step depends only on its last output rather than all the past outputs. This significantly reduces the computational cost of implementing the privacy filter, compared with [26], as the Kalman filtering computations are not required.

C. Contributions

To highlight the vulnerability of the vehicular control systems against inference attacks, this paper investigates an inference attack on the controller gains associated with the lateral velocity, yaw rate and the driver steering command of an active steering control system (ASCS), a control system commonly implemented in modern vehicles to improve the lateral stability. In our set-up, a curious-but-honest adversary receives the lateral velocity, yaw rate and driver steering command of a vehicle via a vehicular ad hoc network (VANET) and attempts to infer the control gains of the ASCS.

To demonstrate the potential privacy breach of controller gains, we use the least squares estimator to infer the gains based on the yaw rate, lateral velocity, and driver steering command. According to our numerical analysis, an adversary with access to this information can reliably infer the controller gains. We next propose a randomized filtering approach for protecting the controller gains against inference attacks. In this approach, the privacy is ensured by means of a filter that consists of two components: a randomizer and a nonlinear transformation.

The randomizer randomly generates a vector of pseudo gains which are different from the true gains of the ASCS system. The nonlinear transformation alters the lateral velocity and yaw rate measurements such that the statistical distribution of the output of the filter is characterized by the pseudo gains rather than the true gains. The nonlinear transformation has a feedforward-feedback structure which enables real-time and causal computation of the pseudo measurements. The output of the nonlinear transformation is then shared via the VANET. In our set-up, the randomization probabilities of the randomizer are designed by solving a convex optimization problem which minimizes the distortion due to the filter subject to a constraint on the privacy of gains. The privacy constraint determines the protection level of the controller gains against inference attacks by imposing an upper bound on the mutual information between the true gains and the pseudo gains. Using information-theoretic inequalities, we show that the

TABLE I
VEHICLE SPECIFICATIONS

Description	Parameter/State	Unit
Vehicle mass	m	[kg]
Vehicle moment of inertia	I_z	[kg.m^2]
Front & rear axles to CG	a, b	[m]
Vehicle lateral velocity at CG	v	[m/s]
Vehicle longitudinal velocity at CG	u	[m/s]
Vehicle yaw rate	r	[rad/s]
Tire cornering stiffness	$C_{\alpha_r}, C_{\alpha_f}$	[Ns/m]

performance of any estimator, that uses the output of the filter to estimate the gains, is limited by the privacy constraint. The performance of the proposed privacy filter is compared with that of an additive noise privacy mechanism. Our numerical results show that the proposed privacy filter significantly outperforms the additive noise mechanism, especially in the low distortion regime.

D. Outline

The rest of this paper is organized as follows. Section II introduces the active steering system. Section III describes the attack model against the controller gains of the active steering control system. The proposed randomized filtering framework is introduced and analyzed in Section IV. The performance of the proposed filtering scheme is numerically investigated for an active steering system in Section V. Section VI concludes the paper.

II. ACTIVE STEERING CONTROL SYSTEM

One of the main aspects of a vehicle's stability is to provide reliable lane keeping. The objective of the lane keeping problem is to control the vehicle such that its center of gravity follows a specified path. More precisely, the vehicle's yaw angle must follow the path heading while there should be no lateral drift from the path. For lateral maneuvers, the control input is the steering wheel angle which is designed based on the yaw rate and lateral velocity, measured by an Inertial Measurement Unit (IMU), and the driver's feed-forward command. The vehicle's lateral offset, measured by a vision system as the distance between the road centerline and a virtual point at a fixed distance from the vehicle, has also been used in the design of control action. However, such a vision-based measurement does not exist in all passenger vehicles.

In this section, we first discuss the steering control for vehicle lane keeping problem. We next discuss the controller gain privacy in the active steering system.

A. Active Steering for Vehicle Lane Keeping

We use a conventional two degrees of freedom bicycle model for vehicle handling dynamics, shown in Fig. 1 (a). Based on this model, the yaw rate and vehicle's lateral velocity evolve based on the following state-space form [27]

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + BM_z(t) + E\boldsymbol{\delta}(t) + \boldsymbol{w}(t), \tag{1}$$

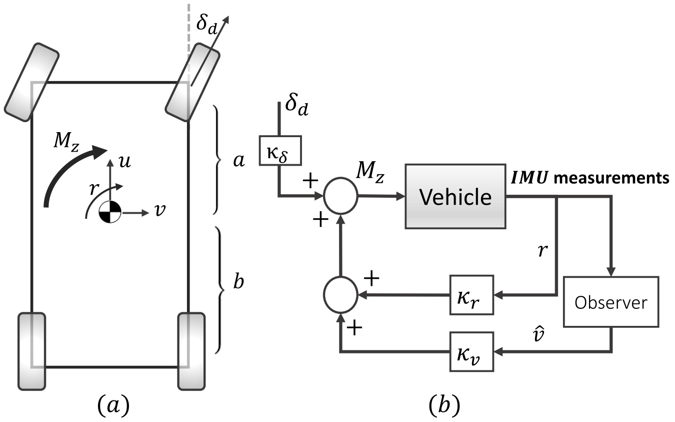

Fig. 1. (a) Plan view of vehicle dynamics model. (b) Vehicle handling control loop.

where $\boldsymbol{x}(t) = [\boldsymbol{v}(t), \boldsymbol{r}(t)]^\top$, $\boldsymbol{v}(t)$ and $\boldsymbol{r}(t)$ are the lateral velocity and yaw rate at time t, respectively, M_z is the yaw moment which acts as the control input, $\boldsymbol{\delta}(t)$ is the driver steering commands, $\boldsymbol{w}(t)$ is the process noise, and

$$A = \begin{bmatrix} -2\frac{C_{\alpha_f}+C_{\alpha_r}}{um} & 2\frac{bC_{\alpha_r}-aC_{\alpha_f}}{um}-u \\ 2\frac{bC_{\alpha_r}-aC_{\alpha_f}}{uI_z} & -2\frac{a^2C_{\alpha_f}+b^2C_{\alpha_r}}{uI_z} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ \frac{1}{I_z} \end{bmatrix}, \quad E = \begin{bmatrix} \frac{2C_{\alpha_f}}{m} \\ \frac{2aC_{\alpha_f}}{I_z} \end{bmatrix}. \tag{2}$$

Table I shows the description of various parameters affecting the evolution of the yaw rate and lateral velocity. The control law consists of a linear combination of the feed-forward driver steering command $\boldsymbol{\delta}(t)$ and the two state-feedback terms based on $\boldsymbol{r}(t)$ and $\boldsymbol{v}(t)$. According to the current state of technology, the direct measurements of the yaw rate $\boldsymbol{r}(t)$ and the steering angle $\boldsymbol{\delta}(t)$ are quite feasible. Direct measurement of the lateral velocity $\boldsymbol{v}(t)$ might be impractical. However, it can be accurately estimated using various estimation techniques, *e.g.,* see [28]. Thus, we use the estimated lateral velocity $\hat{\boldsymbol{v}}(t)$. The control law M_z, as shown in Fig. 1 (b), takes the following form

$$M_z(t) = \kappa_v \boldsymbol{v}(t) + \kappa_r \boldsymbol{r}(t) + \kappa_\delta \boldsymbol{\delta}(t), \tag{3}$$

where κ_v, κ_r, and κ_δ are the lateral velocity feedback gain, the yaw rate feedback gain and the steering angle feed-forward gain, respectively. In (3), we assumed that the lateral velocity can be estimated accurately, *i.e.,* $\hat{\boldsymbol{v}}(t) \approx \boldsymbol{v}(t)$. Substituting (3) into (1) we get

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B[\boldsymbol{\kappa}_v, \boldsymbol{\kappa}_r]\boldsymbol{x}(t) + (\kappa_\delta B + E)\boldsymbol{\delta}(t) + \boldsymbol{w}(t).$$

Discretizing the equation above with the step-size T_s, we obtain the following discrete-time model

$$\boldsymbol{x}_{n+1} = (I + T_s A + T_s B[\boldsymbol{\kappa}_v, \boldsymbol{\kappa}_r])\boldsymbol{x}_n + T_s(\boldsymbol{\kappa}_\delta B + E)\boldsymbol{\delta}_n + \boldsymbol{w}_n, \tag{4}$$

where $n = 1, 2, \ldots,$ is the time index, I is a two-by-two identity matrix, $\boldsymbol{\delta}_n$ is the driver steering command at time-step

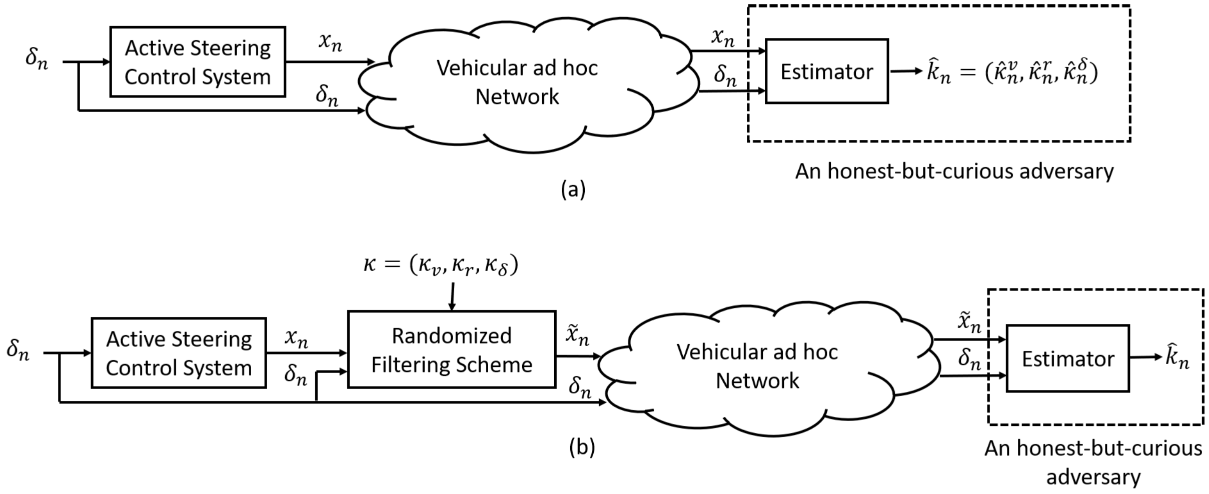
δ_n
Active Steering Control System
x_n
δ_n
Vehicular ad hoc Network
x_n
δ_n
Estimator
$\hat{\kappa}_n = (\hat{\kappa}_n^v, \hat{\kappa}_n^r, \hat{\kappa}_n^\delta)$
An honest-but-curious adversary
(a)
$\kappa = (\kappa_v, \kappa_r, \kappa_\delta)$
δ_n
Active Steering Control System
x_n
δ_n
Randomized Filtering Scheme
$\tilde{x}_n$
δ_n
Vehicular ad hoc Network
$\tilde{x}_n$
δ_n
Estimator
$\hat{\kappa}_n$
An honest-but-curious adversary
(b)

Fig. 2. The information sharing between an ASCS and a vehicular ad hoc network (*a*) without the randomized filtering scheme, (*b*) with the proposed randomized filtering scheme.

n and $\boldsymbol{x}_n = (\boldsymbol{v}_n, \boldsymbol{r}_n)^\top$ is the state of the ASCS at time-step n where $\boldsymbol{v}_n$ and $\boldsymbol{r}_n$ denote the lateral velocity and the yaw rate at time-step n, respectively.

B. Notations and Standing Assumptions

We use $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_v, \boldsymbol{\kappa}_r, \boldsymbol{\kappa}_\delta)$ to denote the controller gains associated with the yaw rate and lateral velocity measurements where $\boldsymbol{\kappa}$ takes values in the set $\mathcal{K} = \{\kappa_1, \ldots, \kappa_m\}$ with probability $\mathsf{Pr}(\boldsymbol{\kappa} = \kappa_i) = p_i$. This assumption is motivated by the fact that the gain scheduling technique is commonly used to design controllers for ASCSs, *e.g.,* see [29] and [30]. In this approach, a set of gains are designed for an ASCS and depending on the operating point of the system one of the gains is implemented. The realization of $\boldsymbol{\kappa}$ is denoted by $\kappa = (\kappa_v, \kappa_r, \kappa_\delta)$. We assume that the gains are fixed over a horizon of length T. The process noise $\{\boldsymbol{w}_n\}$ in (4) is modeled as a sequence of independent and identically distributed (i.i.d.) random variables. The common distribution of $\{\boldsymbol{w}_n\}_n$ is assumed to be absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^2$. Let $\boldsymbol{x}_{1:n}$ denote the sequence of the ASCS's states over the horizon $1, \ldots, n$. A realization of $\boldsymbol{x}_{1:n}$ is denoted by $x_{1:n}$. Let $p_\kappa(x_{1:n}; \delta_{1:n})$ denote the joint probability density function (p.d.f.) of $\boldsymbol{x}_{1:n}$ when the vector of true gains $\boldsymbol{\kappa}$ is equal to κ and $\boldsymbol{\delta}_{1:n} = \delta_{1:n}$. Note that the joint p.d.f. of the states of the ASCS is parameterized by the control gains. Thus, it belongs to the set $\mathcal{M} = \{p_\kappa(x_{1:n}; \delta_{1:n})\}_{\kappa \in \mathbb{R}^3}$.

The conditional cumulative distribution function (c.d.f.) of $\boldsymbol{v}_{n+1}$ given the event $\{\boldsymbol{v}_n = v_n, \boldsymbol{r}_n = r_n, \boldsymbol{\delta}_n = \delta_n, \boldsymbol{\kappa} = \kappa\}$ is denoted by $F_v(\cdot\,|v_n, r_n, \delta_n, \kappa)$ which is defined as

$$F_v(z\,|v_n, r_n, \delta_n, \kappa) = \int_{-\infty}^{z} p_{\kappa,v}(z\,|v_n, r_n, \delta_n)\,dz,$$

where $p_{\kappa,v}(x\,|v_n, r_n, \delta_n)$ is the conditional p.d.f. of $\boldsymbol{v}_{n+1}$ given the event $\{\boldsymbol{v}_n = v_n, \boldsymbol{r}_n = r_n, \boldsymbol{\delta}_n = \delta_n, \boldsymbol{\kappa} = \kappa\}$. Similarly, the conditional c.d.f. of $\boldsymbol{r}_{n+1}$ given $\{\boldsymbol{v}_{n+1} = v_{n+1}, \boldsymbol{v}_n = v_n, \boldsymbol{r}_n = r_n, \boldsymbol{\delta}_n = \delta_n, \boldsymbol{\kappa} = \kappa\}$, is denoted by $F_r(\cdot\,|v_{n+1}, v_n, r_n, \delta_n, \kappa)$ and is defined as

$$F_r(z\,|v_{n+1}, v_n, r_n, \delta_n, \kappa) = \int_{-\infty}^{z} p_{\kappa,r}(z\,|v_{n+1}, v_n, r_n, \delta_n)\,dz,$$

where $p_{\kappa,r}(x\,|v_{n+1}, v_n, r_n, \delta_n)$ is the conditional p.d.f. of $\boldsymbol{r}_{n+1}$ given the event $\{\boldsymbol{v}_{n+1} = v_{n+1}, \boldsymbol{v}_n = v_n, \boldsymbol{r}_n = r_n, \boldsymbol{\delta}_n = \delta_n, \boldsymbol{\kappa} = \kappa\}$. We also follow the convention that

$$F_v(x\,|v_0, r_0, \delta_0, \kappa) = F_v(x),$$
$$F_r(x\,|v_1, v_0, r_0, \delta_0, \kappa) = F_r(x\,|v_1).$$

where $F_v(x)$ is the c.d.f. of the initial lateral velocity $\boldsymbol{v}_1$ and $F_r(x\,|v_1)$ is the conditional c.d.f. of initial yaw rate $\boldsymbol{r}_1$ given the initial lateral velocity $\{\boldsymbol{v}_1 = v_1\}$. In the rest of the paper, we assume that the sequence of random variables $\{\boldsymbol{v}_n = v_n, \boldsymbol{r}_n = r_n, \boldsymbol{\delta}_n = \delta_n\}_n$ has a continuous joint probability distribution which is absolutely continuous with respect to the Lebesgue measure.

III. Inference Attack Model

Consider a vehicular ad hoc network (VANET) wherein a vehicle may exchange information with other vehicles as well as the transportation infrastructure, *e.g.,* connected traffic lights. In a VANET, each vehicle transmits various pieces of information including (but not limited to): temporal ID of the vehicle, time of transmission, latitude, longitude and elevation of the vehicle, longitudinal and lateral velocities and acceleration of the vehicle, yaw rate, and steering and break status of the vehicle [31], [32]. This information is then used to improve safety and efficiency levels of the transportation systems.

In an intelligent transportation system, the computing units of certain vehicles (or infrastructure equipment) might be compromised, *e.g.,* via malware, to act as honest-but-curious adversaries, *i.e.,* the adversarial agents that have lawful access to the shared information by vehicles via VANETs, and may attempt to infer the *parameters of control systems* of vehicles based on the shared data. The privacy breach of control parameters exposes the manufacturers as well as the system designers to severe financial risks due to the sheer monetary values of the design parameters of vehicles. Additionally, such privacy breaches might also be exploited by adversarial agents, *e.g.,* cyber-attackers, to launch detrimental model-based attacks on vehicles [3]–[6].

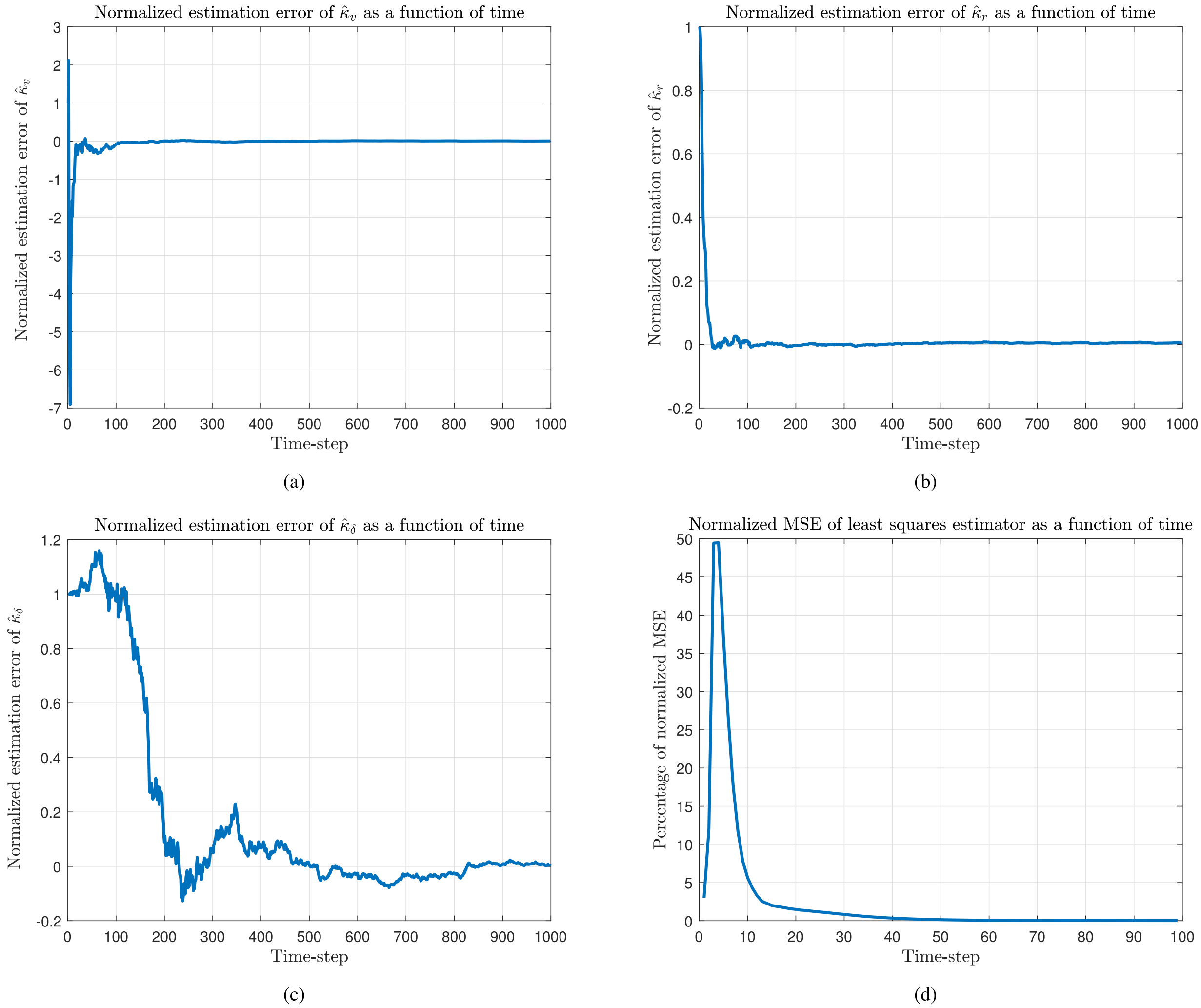

Fig. 3. Normalized estimation error of $\hat{\boldsymbol{\kappa}}_v$ (a), $\hat{\boldsymbol{\kappa}}_r$ (b), and $\hat{\boldsymbol{\kappa}}_\delta$ (c) as a function of time. Percentage of the normalized MSE of the least squares estimator as a function of time (d).

In this paper, we study the privacy filter design problem for protecting the control gains of an ASCS against inference attacks. To this end, we first discuss the attacker model in the next subsection. We then present the numerical results of an inference attack on an ASCS.

A. Attacker Model

In this paper, we model the attacker as an honest-but-curious adversary that attempts to infer the controller gains of the ASCS of a vehicle based on the yaw rate, lateral velocity and the driver steering command of that vehicle, as shown in Fig. 3.(a). Note that yaw rate, lateral velocity, and the driver steering command are continuously shared via VANETs under different vehicular communication standards such as SAE J2735 [31], [32]. Through this paper, we assume that the attacker has knowledge of the control system structure and uses a least squares estimator to infer the control gains of the ASCS.

B. An Inference Attack on an ASCS

In this subsection, we show that the adversary is able to reliably infer the controller gains. To this end, we simulated an ASCS using equation (4) (see Section V for the parameters of the ASCS). A least squares estimator was used to infer the control gains based on the states of the ASCS. Let $\hat{\boldsymbol{\kappa}}_v$, $\hat{\boldsymbol{\kappa}}_r$, and $\hat{\boldsymbol{\kappa}}_\delta$ denote the least squares estimator of the $\boldsymbol{\kappa}_v$, $\boldsymbol{\kappa}_r$, and $\boldsymbol{\kappa}_\delta$, respectively. Fig. 3(a)-3(c) show the normalized estimation error of $\hat{\boldsymbol{\kappa}}_v$, $\hat{\boldsymbol{\kappa}}_r$, and $\hat{\boldsymbol{\kappa}}_\delta$ as a function time for a realization of $(\boldsymbol{x}_{1:T}, \boldsymbol{\delta}_{1:T})$. The normalized estimation error is defined as the ratio of the estimation error to the true gain. As these figures show, the least squares estimator is able to accurately identify the values of controller gains when the states of the ASCS are directly shared via a VANET.

Fig. 3(d) shows the percentage of the normalized means square error (MSE) of the least squares estimator as a function of the number of samples where the normalized MSE is defined as

$$\mathsf{E}\left[\left(\frac{\hat{\boldsymbol{\kappa}}_v - \boldsymbol{\kappa}_v}{\boldsymbol{\kappa}_v}\right)^2 + \left(\frac{\hat{\boldsymbol{\kappa}}_r - \boldsymbol{\kappa}_r}{\boldsymbol{\kappa}_r}\right)^2 + \left(\frac{\hat{\boldsymbol{\kappa}}_\delta - \boldsymbol{\kappa}_\delta}{\boldsymbol{\kappa}_\delta}\right)^2\right].$$

The normalized MSE is computed using 1000 realizations of $(\boldsymbol{x}_{1:T}, \boldsymbol{\delta}_{1:T})$. Based on this figure, the least squares estimator can reliably infer the controller gains for all the considered realizations. These observations indicate that directly sharing

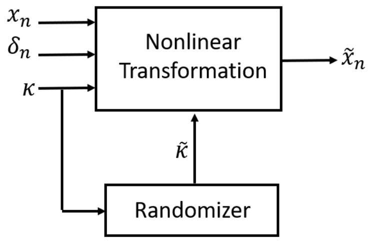

Fig. 4. The structure of the proposed randomized filtering scheme.

the states of the ASCS via a VANET exposes the controller gains to the risk of inference attacks.

IV. The Randomized Filtering Framework

In Section III, we demonstrated that an adversary with access to the states of the ASCS ($\boldsymbol{x}_{1:T}$) and the driver steering command ($\boldsymbol{\delta}_{1:T}$) can reliably infer the controller gains. To overcome this problem, in this section, we propose a randomized filtering framework to ensure that the controller gains cannot be reliably estimated based on the shared information. In the proposed framework, a filter at each time-step n takes $\boldsymbol{x}_n$, $\boldsymbol{\delta}_n$ and $\boldsymbol{\kappa}$ as input, and generates an output denoted by $\tilde{\boldsymbol{x}}_n$, as shown in Fig. 2(b). Then, the output of the filter ($\tilde{\boldsymbol{x}}_n$) and $\boldsymbol{\delta}_n$ are shared via a vehicular ah hoc network (VANET).

The filter is designed to achieve the following three objectives:

1) The filter's output should accurately represent the states of the ASCS.
2) The output of the filter should not be an informative source for estimating the controller gains.
3) The statistical distribution of the states of the ASCS ($\boldsymbol{x}_{1:T}$) and that of the filter's output ($\tilde{\boldsymbol{x}}_{1:T}$) should belong to the same family of distributions.

Note that, given $\boldsymbol{\kappa} = \kappa$, the joint probability density function (p.d.f.) of $\boldsymbol{x}_{1:T}$ belongs to the family of density functions $\mathcal{M} = \{p_\kappa(x_{1:T}; \delta_{1:T})\}_{\kappa \in \mathbb{R}^3}$. The objective 3 guarantees that the joint p.d.f. of the filter's output over the horizon $1, \ldots, T$ also belongs to $\mathcal{M}$. Hence, the proposed randomized filtering scheme preserves the structure of the statistical model of the states of the ASCS.

Remark 1: In a VANET, the shared information by a vehicle is used in the control, prediction and estimation algorithms in other vehicles. The complexity of these algorithms depends on the complexity of the model of the shared information, *i.e.,* a highly complex model demands complicated signal processing and control algorithms. Without objective 3, the statistical model of the filter's output might become overly complex which results in highly complicated algorithms. □

To achieve these objectives 1-3, a randomized filtering framework is proposed which consists of two components: a nonlinear transformation and a randomizer as shown in Fig. 4. The randomizer takes the true controller gains ($\boldsymbol{\kappa} = (\boldsymbol{\kappa}_v, \boldsymbol{\kappa}_r, \boldsymbol{\kappa}_\delta)$) as input and randomly generates the vector of *pseudo gains* $\tilde{\boldsymbol{\kappa}} = \left(\tilde{\boldsymbol{\kappa}}_v, \tilde{\boldsymbol{\kappa}}_r, \tilde{\boldsymbol{\kappa}}_\delta\right)$ which takes values in the set $\tilde{\mathcal{K}} = \{\tilde{\kappa}_1, \ldots, \tilde{\kappa}_{\tilde{m}}\}$. At each time-step n, the nonlinear transformation takes $\boldsymbol{\kappa}$, $\tilde{\boldsymbol{\kappa}}$, $\boldsymbol{x}_n$ and $\boldsymbol{\delta}_n$ as input. Then, it generates the vector of pseudo states $\tilde{\boldsymbol{x}}_n = \left(\tilde{\boldsymbol{v}}_n, \tilde{\boldsymbol{r}}_n\right)^\top$ where $\tilde{\boldsymbol{v}}_n$ and $\tilde{\boldsymbol{r}}_n$ are the pseudo lateral velocity and the pseudo yaw rate, respectively.

In the remainder of this section, we will first discuss the structure of the nonlinear transformation followed by the optimal design of the randomizer and the privacy analysis of the controller gains under the proposed framework.

A. Nonlinear Transformation

In this subsection, we will describe the objective and the structure of the nonlinear transformation. To this end, suppose that the vector of true controller gains is equal to κ_i and the randomizer has selected $\tilde{\kappa}_j$ as the vector of pseudo gains. Thus, the joint p.d.f. of the true measurements is given by $p_{\kappa_i}(x_{1:T}; \delta_{1:T})$. The nonlinear transformation ensures that the joint p.d.f. of the output of the filter over the horizon $1, \ldots, T$ is given by $p_{\tilde{\kappa}_j}(\tilde{x}_{1:T}; \delta_{1:T})$.

Let $\tilde{x}_n = \left[\tilde{v}_n, \tilde{r}_n\right]^\top$ denote a realization of the output of the filter at time-step n. Then, given $\boldsymbol{\kappa} = \kappa_i$ and $\tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j$, the filter, at time-step n, first generates the pseudo lateral velocity according to

$$\tilde{v}_n = F_v^{-1}\left(d_n^1 \left| \tilde{v}_{n-1}, \tilde{r}_{n-1}, \delta_{n-1}, \tilde{\kappa}_j \right.\right), \tag{5}$$

where d_n^1 is given by

$$d_n^1 = F_v\left(v_n \left| v_{n-1}, r_{n-1}, \delta_{n-1}, \kappa_i \right.\right),$$

Next, the filter generates the pseudo yaw rate according to

$$\tilde{r}_n = F_r^{-1}\left(d_n^2 \left| \tilde{v}_n, \tilde{v}_{n-1}, \tilde{r}_{n-1}, \tilde{\kappa}_j \right.\right), \tag{6}$$

where d_n^2 is given by

$$d_n^2 = F_r\left(r_n \left| v_n, v_{n-1}, r_{n-1}, \kappa_i \right.\right).$$

The structure of the nonlinear transformation is shown in Fig. 5. Next theorem studies the statistical distribution of the output of the filter over the horizon $1, \ldots, T$.

Theorem 1: Consider the nonlinear transformation specified by equations (5) and (6). Given $\boldsymbol{\kappa} = \kappa_i$ and $\tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j$, the joint probability density function of the pseudo states $\tilde{\boldsymbol{x}}_{1:T}$ is given by $p_{\tilde{\kappa}_j}(\tilde{x}_{1:T}, \delta_{1:T})$, for all i, j. □

Proof: See Appendix A. ■

According to Theorem 1, the nonlinear transformation ensures that the joint p.d.f. of the output of the filter is parameterized by the vector of pseudo gains $\tilde{\boldsymbol{\kappa}}$ rather than the true gains. Theorem 1 also implies that, given $\tilde{\boldsymbol{\kappa}}$, the joint p.d.f. of the output of the filter belongs to the family of density functions $\mathcal{M} = \{p_\kappa(x_{1:T}; \delta_{1:T})\}_{\kappa \in \mathbb{R}^3}$. Hence, the proposed transformation ensures that the input and output of the filter belong to the same family of distribution functions (objective 3 of Section IV).

A unique aspect of the proposed privacy is the *feedforward-feedback* structure of the nonlinear transformation as shown in in Fig. 5. The feedforward component of the filter computes d_n^1 and d_n^2 whereas the feedback component computes $\tilde{v}_n$ and $\tilde{r}_n$ using the last output of the filter. Moreover, the recursive structure of the proposed privacy filter allows causal (real-time) generation of the pseudo measurements. Thus, $\tilde{v}_n$ and

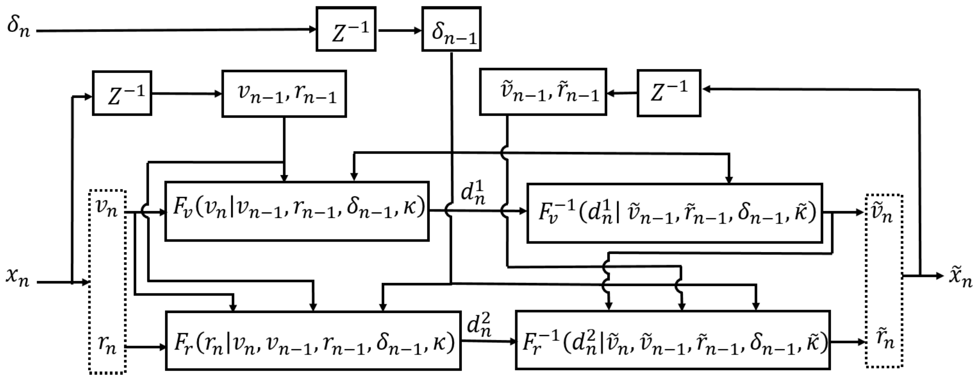
δ_n
Z^{-1}
δ_{n-1}
Z^{-1}
v_{n-1}, r_{n-1}
$\tilde{v}_{n-1}, \tilde{r}_{n-1}$
Z^{-1}
v_n
$F_v(v_n|v_{n-1}, r_{n-1}, \delta_{n-1}, \kappa)$
d_n^1
$F_v^{-1}(d_n^1|\,\tilde{v}_{n-1}, \tilde{r}_{n-1}, \delta_{n-1}, \tilde{\kappa})$
$\tilde{v}_n$
x_n
$\tilde{x}_n$
r_n
$F_r(r_n|v_n, v_{n-1}, r_{n-1}, \delta_{n-1}, \kappa)$
d_n^2
$F_r^{-1}(d_n^2|\tilde{v}_n, \tilde{v}_{n-1}, \tilde{r}_{n-1}, \delta_{n-1}, \tilde{\kappa})$
$\tilde{r}_n$

Fig. 5. The structure of the nonlinear transformation at time-step n.

$\tilde{r}_n$ are causally generated using the sensor measurements up to time n rather than $r_1, v_1, \ldots, r_T, v_T$.

B. Randomizer

The randomizer generates the vector of pseudo gains $\tilde{\boldsymbol{\kappa}}$ based on the true values of gains according to the stochastic kernel $\pi\left(\cdot\,|\cdot\right)$:

$$\pi\left(\tilde{\kappa}_j\,|\kappa_i\right) = \mathsf{Pr}\left(\tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j \middle| \boldsymbol{\kappa} = \kappa_i\right), \quad \forall i, j.$$

That is, $\pi\left(\tilde{\kappa}_j\,|\kappa_i\right)$ specifies the probability that the randomizer selects $\tilde{\kappa}_j$ as the vector of pseudo gains when the vector of true gains is equal to κ_i. The randomization probabilities $\left\{\pi\left(\tilde{\kappa}_j\,|\kappa_i\right)\right\}_{i,j}$ are the design parameters of the randomizer. To discuss the optimal choice of the randomization probabilities, we define the total average distortion between the true and pseudo states over the horizon $1, \ldots, T$ as

$$\frac{1}{T}\sum_{n=1}^{T}\mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_x\right\|^2\right].$$

Note that due to the filter, the true yaw rate and lateral velocity might be different from the pseudo yaw rate and lateral velocity. Thus, the total average distortion captures the average deviation of the output of the filter from its input.

The optimal randomization probabilities are obtained by minimizing the total average distortion subject to a constraint on the privacy of controller gains that determines the protection level of the controller gains against inference attacks. In this paper, we use the mutual information between the true gains and the pseudo gains as the privacy metric which is defined as

$$\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right] = \sum_{i,j}\mathsf{Pr}\left(\boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right)\log\frac{\mathsf{Pr}\left(\boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right)}{\mathsf{Pr}\left(\boldsymbol{\kappa} = \kappa_i\right)\mathsf{Pr}\left(\tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right)}.$$

This privacy metric captures the amount of information that can be inferred about the true gains by observing pseudo gains. A small value of $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]$ indicates a low level of leakage of private information as the true gains cannot be reliably inferred using the pseudo gains when $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]$ is small. A large value of $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]$ implies a large level of information leakage. Hence, on can reliably infer the true gains by observing the pseudo gains when $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]$ is large. Thus, a low level of $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]$ is desirable for ensuring privacy.

The optimal randomization probabilities are the solution of the following optimization problem:

$$\begin{aligned}
\underset{\left\{\pi\left(\tilde{\kappa}_j|\kappa_i\right)\right\}_{i,j}}{\text{minimize}} \quad & \frac{1}{T}\sum_{n=1}^{T}\mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_n\right\|^2\right] \\
& \pi\left(\tilde{\kappa}_j\,|\kappa_i\right) \geq 0, \quad \forall i, j \\
& \sum_j \pi\left(\tilde{\kappa}_j\,|\kappa_i\right) = 1 \quad \forall i \\
& \mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right] \leq I_0,
\end{aligned} \tag{7}$$

where the second constraint ensures the law of total probability and the last constraint imposes an upper bound on the mutual information between the input and the output of the randomizer. We refer to the last constraint in (7) as the *privacy constraint* since it limits the amount of information that can be inferred about the true gains by observing the pseudo gains. We also refer to I_0 as the *level of information leakage*. In the next subsection, we show that the privacy constraint limits the ability of any adversary in estimating the true gains based on the output of the filter. Next theorem studies the structure of the optimization problem (7).

Theorem 2: The objective function in the optimization problem (7) is linear in the randomization probabilities. Moreover, the privacy constraint is convex. □

Proof: See Appendix B. ■

According to Theorem 2, the optimization problem (7) is convex. Thus, the optimal randomization probabilities can be obtained by solving a convex optimization problem using efficient numerical techniques. Note that the optimal design of randomization probabilities ensures that the total distortion is minimized while a certain privacy level for the controller gains is guaranteed. Thus, the proposed filtering scheme achieves the objectives 1 and 2 in Section IV.

C. Privacy Level of Control Gains

In this subsection, we study the privacy level of controller gains under the proposed filtering scheme. To this end, consider an adversary with access to $\tilde{\boldsymbol{x}}_{1:T}$, $\boldsymbol{\delta}_{1:T}$ and $\mathcal{K}$. The adversary is interested in inferring the controller gains employed by the ASCS over the horizon $1,\ldots,T$. With an abuse of notation, let $\hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)$ denote an arbitrary estimator of the true gains based on the output of the filter where the estimator is defined as a mapping from $\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}$ to the set $\mathcal{K}$. To evaluate the privacy level of gains under the proposed framework, next theorem establishes a lower bound on the error probability of the estimator $\hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)$.

Theorem 3: Let $\mathsf{Pr}\left(\boldsymbol{\kappa}\neq\hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)\right)$ denote the error probability of the estimator of the controller gains $\hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)$. Then, we have

$$\mathsf{Pr}\left(\boldsymbol{\kappa}\neq\hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)\right)\geq\frac{\mathsf{H}\left[\boldsymbol{\kappa}\right]-I_0-1}{\log|\mathcal{K}|}. \tag{8}$$

where $|\mathcal{K}|$ is the cardinality of $\mathcal{K}$, I_0 is the upper bound on the privacy constraint in (7) and $\mathsf{H}\left[\boldsymbol{\kappa}\right]$ is the discrete entropy of $\boldsymbol{\kappa}$. □

Proof: See Appendix C. ■

Theorem 3 establishes a lower bound on the performance of any estimator that uses the output of the filter over the horizon $1,\ldots,T$ to recover the true gains employed by the controller. This lower bound depends on the discrete entropy of $\boldsymbol{\kappa}$, I_0 and the cardinality of the set of gains. According to Theorem 3, the lower bound on the error probability of any estimator of gains increases as I_0 becomes small. Thus, a small value of I_0 ensures that even an adversary with access to the set of controller gains $\mathcal{K}$ cannot reliably estimate the true gains employed by the ASCS over the horizon $1,\ldots,T$.

In Appendix C, we show that the mutual information between the true gains and the shared information can be upper bounded by the mutual information between the true and pseudo gains. That is, we have

$$\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right]\leq\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{\kappa}}\right]. \tag{9}$$

Note that $\mathsf{I}\left[\boldsymbol{\kappa};\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right]$ quantifies the amount of information which can be inferred about the true gains based on the output of the filter. Thus, the inequality (9) implies that the privacy constraint in the optimization problem (7) essentially limits the leakage of information about the true gains via the filter's output.

Remark 2: To prove Theorem 3, we establish the Markov chain $\boldsymbol{\kappa}\rightarrow\left(\tilde{\boldsymbol{\kappa}},\boldsymbol{\delta}_{1:T}\right)\rightarrow\tilde{\boldsymbol{x}}_{\mathbf{1:T}}$ in Appendix A. Then, the data processing inequality [33] and this Markov chain are used to derive the inequality (9). Finally, the lower bound in Theorem 3 is established using Fano's inequality and the inequality (9). □

Remark 3: An attacker with access to the filter's output can reliably infer the pseudo gains. If the true gains of the system and the pseudo gains are close, then the attacker obtains a good estimate of the true gains by inferring the pseudo gains. However, when the difference between the true gains and pseudo gains is large, the knowledge of pseudo gains will not provide a good estimate of the true gains. During the design process of the privacy filter, the designer can ensure that the attacker cannot obtain an accurate estimate of the true gains by inferring the pseudo gains. This objective can be achieved by designing the set of pseudo gains such that pseudo gains and true gains are far enough from each other. □

TABLE II

PARAMETERS OF THE ACTIVE STEERING SYSTEM

Parameter	Value
a	1.25 m
b	1.2 m
m	1300 Kg
C_{α_f}	15000 N/rad
C_{α_r}	15000 N/rad
u	20, 40, 50 m/s
I_z	2500 kgm^2
T_s	0.1s
T	1000 samples

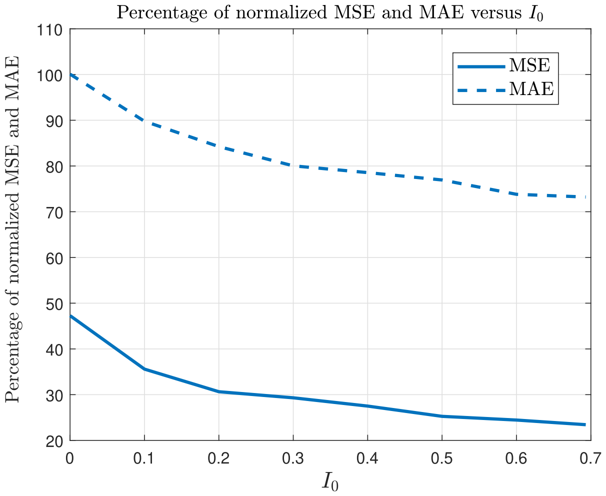

Fig. 6. The percentage of the normalized MSE and MAE error of the gain estimator versus the level of information leakage I_0.

V. NUMERICAL RESULTS

In this section, we will numerically investigate the performance of the proposed framework in ensuring the privacy of the controller gains of an ASCS. To this end, we consider an active steering system with the dynamics in (4) where the disturbance $\{\boldsymbol{w}_n\}_n$ is a sequence of independent and identically distributed Gaussian random vectors. For each time-step n, the entries of $\boldsymbol{w}_n=\left(\boldsymbol{w}_n^1,\boldsymbol{w}_n^2\right)^\top$ are assumed to be independent Gaussian random variables with zero mean and variance equal to 10^{-2}. In our set-up, the vector of true controller gains $\boldsymbol{\kappa}=(\boldsymbol{\kappa}_v,\boldsymbol{\kappa}_r,\boldsymbol{\kappa}_\delta)^\top$ takes values in the set

$$\mathcal{K}=\left\{\begin{bmatrix}777.9\\39727\\10000\end{bmatrix},\begin{bmatrix}477.9\\29727\\15000\end{bmatrix}\right\}$$

with equal probabilities. The vector of pseudo gains $\tilde{\boldsymbol{\kappa}}$ takes values in

$$\tilde{\mathcal{K}}=\left\{\begin{bmatrix}677.9\\34727\\5000\end{bmatrix},\begin{bmatrix}377.9\\24727\\10000\end{bmatrix}\right\}$$

The efficiency of the proposed privacy mechanism is examined by constructing a least squares estimator of controller gains based on $\left(\tilde{\boldsymbol{x}}_{1:T},\boldsymbol{\delta}_{1:T}\right)$. To estimate the controller gains, we generated the pseudo states $\tilde{\boldsymbol{x}}_{1:T}$ for longitudinal velocities of 20, 40, and 50. Then, the pseudo states along with the driver steering commands were used as input to the estimator.

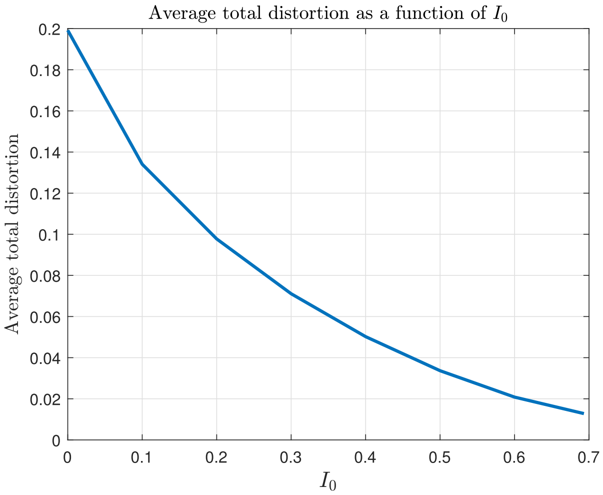

Fig. 7. Average total distortion versus the level of information leakage I_0.

The horizon length was set to 1000 samples. We then computed the normalized means square error (MSE) defined as

$$\mathsf{E}\left[\left(\frac{\hat{\kappa}_v-\kappa_v}{\kappa_v}\right)^2+\left(\frac{\hat{\kappa}_r-\kappa_r}{\kappa_r}\right)^2+\left(\frac{\hat{\kappa}_\delta-\kappa_\delta}{\kappa_\delta}\right)^2\right],$$

where $\hat{\kappa}_v$, $\hat{\kappa}_r$ and $\hat{\kappa}_\delta$ are the least square estimates of the gains associated with the lateral velocity, yaw rate and the driver steering commands. Table II shows the parameters of the ASCS in our numerical analysis.

Fig. 6 shows the percentage of the normalized mean square error (MSE) and mean absolute error (MAE) of the least squares estimator for different values of the level of information leakage I_0. According to this figure, the performance of the least squares deteriorates as I_0 becomes small. This observation confirms that, under the proposed framework, the estimator cannot reliably estimate the controller gains. This is due to the fact that, under the proposed scheme, the joint p.d.f. of $\tilde{\boldsymbol{y}}_{1:T}$ is characterized by the vector of pseudo gains which are different from the true gains.

Fig. 7 shows the average total distortion between the true and pseudo states of the ASCS as a function of I_0. According to this figure, the average total distortion increases as I_0 becomes small. Note that the privacy constraint in (7) becomes tight as I_0 decreases. This results in higher values of distortion since the feasible set of the optimization problem (7) shrinks as I_0 becomes small.

We next compare the performance of the proposed privacy filter with that of an additive noise privacy mechanism wherein random perturbation is added to the sensor measurements to ensure privacy. We note that the additive noise schemes have been extensively studied in the context of data privacy, *e.g.*, see [34] and references therein. To this end, we consider an additive noise mechanism in which Gaussian noise is added to the sensor measurements to ensure privacy. Let $\tilde{x}_k^{\text{add}}$ denote the output of the additive noise mechanism which can be described as

$$\tilde{\boldsymbol{x}}_n^{\text{add}}=\boldsymbol{x}_n+\boldsymbol{N}_n,$$

where $\{\boldsymbol{N}_n\}_n$ is a sequence of i.i.d. Gaussian random vectors with zero mean and covariance matrix σ^2I, and I is a 2-by-2 identity matrix. Under the additive noise mechanism, at time-step n, $\tilde{\boldsymbol{x}}_n$ is shared with the vehicular ad hoc network and an adversary with access to $\{\tilde{\boldsymbol{x}}_n\}_n$ may attempt to infer the true values of control gains.

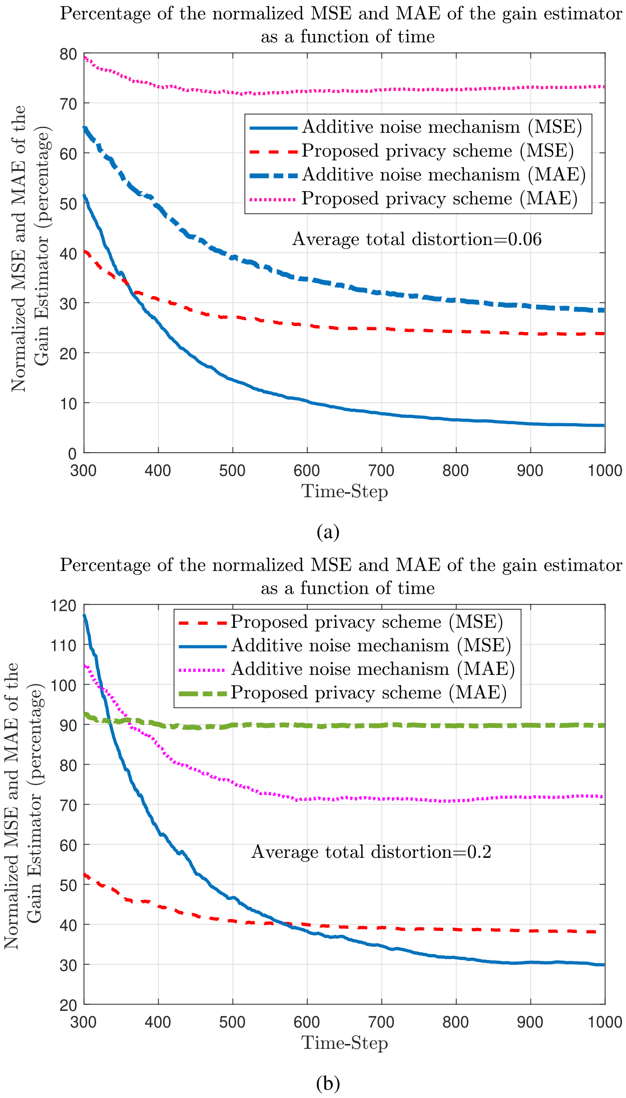

Fig. 8. Percentage of the normalized MSE and MAE of the gain estimator as a function of time when the average total distortion is 0.06 (a), and when the average total distortion is 0.2 (b).

To compare the proposed privacy filter with the additive noise mechanism, we assume that the adversary is aware of the structure of the employed privacy scheme, and uses a least squares estimator to infer the control gains. Thus, in each case, the least squares estimator is designed according to the structure of the privacy mechanism in that case.

Fig. 8(a) illustrates the percentage of the normalized MSE and MAE of the estimator of control gains as a function of time under the proposed privacy filter and the additive noise mechanism when the average total distortion due to each scheme is equal to 0.06. Note that the values of I_0 and σ^2 can be varied to achieve the same distortion level in both cases.

According to Fig. 8(a), the normalized MSE and MAE of the estimator under the additive noise mechanism is 5% when the adversary has access to 1000 measurements. However,

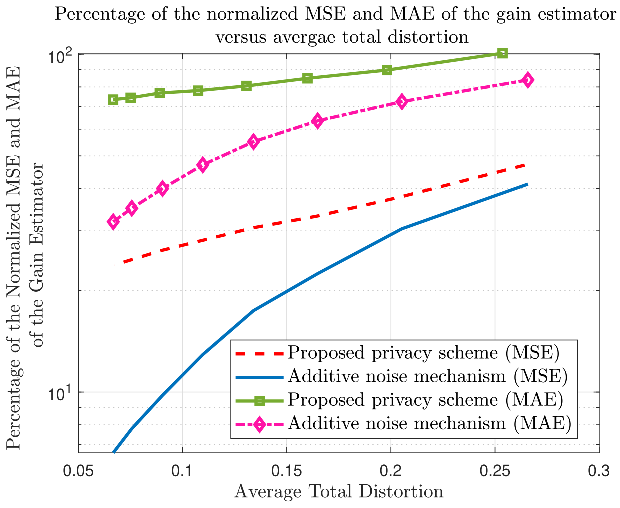

Fig. 9. The percentage of the normalized MSE and MAE error of the gain estimator as a function of the average total distortion under the proposed privacy and the additive noise mechanism.

under the proposed privacy scheme, the normalized MSE of the estimator is close to 25%. This observation indicates that the proposed privacy filter provides a better privacy protection compared with the additive noise mechanism. A same behavior continues to hold for the MAE of the gain estimator. Note that the adversary can accurately estimate the feedback gains under the additive noise mechanism. This is due to the fact that the impact of the additive noise can be averaged out. However, under the proposed privacy filter, the adversary can at most infer pseudo gains accurately and accurate knowledge of pseudo gains does not result in an accurate estimate of true gains. A similar behavior is observed when the average total distortion is 0.2 as shown in Fig. 8(b).

Fig. 9 shows the normalized MSE and MAE of the gain estimator versus the total average distortion under the proposed privacy filter and the additive noise mechanism. The MSE and MAE in each case are computed using 1000 samples of the output of the privacy filter in that case. According to this figure, for a given distortion level, normalized MSE (MAE) of the least squares estimator is higher when the proposed privacy filter in employed compared with that under the additive noise mechanism. This observation also confirms that the proposed privacy scheme is more efficient in ensuring privacy than the additive noise scheme.

VI. Conclusion

In this paper, we proposed a randomized filtering framework for protecting the controller gains of the active steering control system (ASCS) of a vehicle against inference attacks. The proposed framework consists of a randomizer and a nonlinear transformation. The randomizer takes the vector of true gains as input and randomly selects a vector of pseudo gains. The nonlinear transformation takes the true gains, the pseudo gains, the driver's steering command and the states of the ASCS as input. It then generates a vector of pseudo states which is shared with other vehicles and transportation infrastructure via a vehicular ad hoc network. We showed that the randomizer can be optimally designed by solving a convex optimization problem. We also showed that the proposed filtering scheme limits the performance of any estimator in recovering the true gains.

Our results can be extended in multiple directions. The privacy filter design problem under other metrics such as Fisher information and Akaike information criterion is an important and interesting research direction. Another important research direction is the development of computationally efficient methods for the optimal design of the pseudo gains.

Appendix A
Proof of Theorem 1

Recall that $\mathcal{M} = \{p_\kappa(x_{1:T};\delta_{1:T})\}_{\kappa\in\mathbb{R}^3}$ where $p_\kappa(x_{1:T};\delta_{1:T})$ denotes the joint probability density function (p.d.f.) of $\boldsymbol{x}_{1:T}$ when $\boldsymbol{\kappa}=\kappa$ and $\boldsymbol{\delta}_{1:T}=\delta_{1:T}$. To prove Theorem 1, we first derive an expression for $p_\kappa(x_{1:T};\delta_{1:T})$. We then use this expression to show that the joint p.d.f. of the output of the filter, *i.e.,* $\tilde{\boldsymbol{x}}_{1:T}$, also belongs to $\mathcal{M}$. Note that $\boldsymbol{x}_{1:T}=\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{\boldsymbol{T}}$ and $\boldsymbol{x}_n=(\boldsymbol{v}_n,\boldsymbol{r}_n)^\top$. Using the Bayes' rule and the Markov property of the dynamics of the ASCS, $p_\kappa(x_{1:T};\delta_{1:T})$ can be factorized as

$$
\begin{aligned}
&p_\kappa(x_{1:T};\delta_{1:T})\\
&=\prod_{n=0}^{T-1} p_{\kappa,v}\left(v_{n+1}\left|v_n,r_n,\delta_n\right.\right)p_{\kappa,r}\left(r_{n+1}\left|v_{n+1},v_n,r_n,\delta_n\right.\right).
\end{aligned}
$$

We next show that the joint p.d.f. of the output of the filter has the same from as $p_\kappa(x_{1:T};\delta_{1:T})$.

Let $p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right.\right)$ denote the joint p.d.f. of $\tilde{\boldsymbol{x}}_{1:T}$ given $\boldsymbol{\delta}_{1:T}=\delta_{1:T}$, $\boldsymbol{\kappa}=\kappa_i$ and $\tilde{\boldsymbol{\kappa}}=\tilde{\kappa}_j$. We obtain an expression for $p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right.\right)$ as follows. Using the Bayes' rule, we have

$$
\begin{aligned}
&p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right.\right)\\
&\quad=\prod_{n=0}^{T-1}p_{\tilde{\boldsymbol{v}}_{n+1}}\left(\tilde{v}_{n+1}\middle|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)\\
&\quad\quad\times p_{\tilde{\boldsymbol{r}}_{n+1}}\left(\tilde{r}_{n+1}\middle|\tilde{v}_{n+1},\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right),
\end{aligned}\tag{10}
$$

where $p_{\tilde{\boldsymbol{v}}_{n+1}}\left(\cdot\middle|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)$ is the conditional p.d.f. of $\tilde{\boldsymbol{v}}_{n+1}$ given $\left\{\tilde{\boldsymbol{x}}_{1:n}=\tilde{x}_{1:n},\boldsymbol{\kappa}=\kappa_i,\tilde{\boldsymbol{\kappa}}=\tilde{\kappa}_j,\boldsymbol{\delta}_{1:T}=\delta_{1:T}\right\}$ and $p_{\tilde{\boldsymbol{r}}_{n+1}}\left(\cdot\middle|\tilde{v}_{n+1},\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)$ is the conditional p.d.f. of $\tilde{\boldsymbol{r}}_{n+1}$ given $\left\{\tilde{\boldsymbol{v}}_{n+1}=\tilde{v}_{n+1},\tilde{\boldsymbol{x}}_{1:n}=\tilde{x}_{1:n},\boldsymbol{\kappa}=\kappa_i,\tilde{\boldsymbol{\kappa}}=\tilde{\kappa}_j,\boldsymbol{\delta}_{1:T}=\delta_{1:T}\right\}$. We next derive expressions for the conditional p.d.f.s $p_{\tilde{\boldsymbol{v}}_{n+1}}\left(\tilde{v}_{n+1}\left|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right.\right)$. Note that the conditional probability of the event $\left\{\tilde{\boldsymbol{v}}_{n+1}\le y\right\}$ given the event $\left\{\tilde{\boldsymbol{x}}_{\boldsymbol{1:n}}=\tilde{x}_{1:n},\boldsymbol{\kappa}=\kappa_i,\tilde{\boldsymbol{\kappa}}=\tilde{\kappa}_j,\boldsymbol{\delta}_{1:T}=\delta_{1:T}\right\}$ can be written as

$$
\begin{aligned}
&\Pr\left(\tilde{\boldsymbol{v}}_{n+1}\le y\middle|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)\\
&=\Pr\left(F_v^{-1}\left(\boldsymbol{d}_{n+1}^1\left|\tilde{v}_n,\tilde{r}_n,\delta_n,\tilde{\kappa}_j\right.\right)\le y\middle|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)\\
&\overset{(a)}{=}\Pr\left(\boldsymbol{d}_{n+1}^1\le F_v\left(y\left|\tilde{v}_n,\tilde{r}_n,\delta_n,\tilde{\kappa}_j\right.\right)\middle|\tilde{x}_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right)\\
&\overset{(b)}{=}\Pr\left(\boldsymbol{d}_{n+1}^1\le F_v\left(y\left|\tilde{v}_n,\tilde{r}_n,\delta_n,\tilde{\kappa}_j\right.\right)\middle|x_{1:n},\kappa_i,\tilde{\kappa}_j,\delta_{1:T}\right),
\end{aligned}\tag{11}
$$

where (a) holds since $F_v\left(\cdot\left|\tilde{v}_n,\tilde{r}_n,\delta_n,\tilde{\kappa}_j\right.\right)$ is monotonically increasing and continuous, and (b) follows from the fact that, given $\boldsymbol{\kappa}$ and $\tilde{\boldsymbol{\kappa}}$, the transformation from $x_{1:n}$ to $\tilde{x}_{1:n}$ is invertible, hence it is possible to uniquely recover $x_{1:n}$ from $\tilde{x}_{1:n}$ when $\boldsymbol{\kappa}$ and $\tilde{\boldsymbol{\kappa}}$ are known.

We next show that the conditional p.d.f. of d^1_{n+1} given $\left\{\boldsymbol{x}_{1:n} = x_{1:n}, \boldsymbol{\delta}_{1:T} = \delta_{1:T}, \boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right\}$ is a uniform distribution in the interval $[0, 1]$. Note that $F_\upsilon\left(\cdot \left|\tilde{\upsilon}_n, \tilde{r}_n, \delta_n, \tilde{\kappa}_j\right.\right)$ takes values in $[0, 1]$. Also, for $0 \leq y \leq 1$, the conditional probability of the event $\left\{\boldsymbol{d}^1_{n+1} \leq y\right\}$ given the event $\left\{\boldsymbol{x}_{1:n} = x_{1:n}, \boldsymbol{\delta}_{1:T} = \delta_{1:T}, \boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right\}$ can be written as

$$\begin{aligned}
&\Pr\left(d^1_{n+1} \leq y \middle| x_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right)\\
&= \Pr\left(F_\upsilon\left(\boldsymbol{v}_{n+1} \left|\upsilon_n, r_n, \delta_n, \kappa_i\right.\right) \leq y \right| \left. x_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right)\\
&\overset{(a)}{=} \Pr\left(\boldsymbol{v}_{n+1} \leq F_\upsilon^{-1}\left(y|\upsilon_n, r_n, \delta_n, \kappa_i\right) \middle| x_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right)\\
&= \Pr\left(\boldsymbol{v}_{n+1} \leq F_\upsilon^{-1}\left(y|\upsilon_n, r_n, \delta_n, \kappa_i\right) \middle| \upsilon_n, r_n, x_{1:n-1}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right)\\
&\overset{(b)}{=} \Pr\left(\boldsymbol{v}_{n+1} \leq F_\upsilon^{-1}\left(y\left|\upsilon_n, r_n, \delta_n, \kappa_i\right.\right) \middle| \upsilon_n, r_n, \kappa_i, \delta_n\right)\\
&= F_\upsilon\left(F_\upsilon^{-1}\left(y\left|\upsilon_n, r_n, \delta_n, \kappa_i\right.\right)\left|\upsilon_n, r_n, \delta_n, \kappa_i\right.\right)\\
&= y,
\end{aligned} \tag{12}$$

where (a) holds since $F_\upsilon\left(\cdot\left|\upsilon_n, r_n, \kappa_i, \delta_n\right.\right)$ is invertible and (b) follows from the Markov chains $\tilde{\boldsymbol{\kappa}} \rightarrow (\boldsymbol{\kappa}, \boldsymbol{x}_{1:n}, \boldsymbol{\delta}_{1:T}) \rightarrow \boldsymbol{v}_{n+1}$ and $(\boldsymbol{x}_{1:n-1}, \boldsymbol{\delta}_{1:T}) \rightarrow (\boldsymbol{\kappa}, \boldsymbol{x}_n, \boldsymbol{\delta}_n) \rightarrow \boldsymbol{v}_{n+1}$. Equation (12) implies that the random variable d^1_{n+1} given $\left\{\boldsymbol{x}_{1:n} = x_{1:n}, \boldsymbol{\delta}_{1:T} = \delta_{1:T}, \boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right\}$ is uniformly distributed in $[0, 1]$. Combining (12) and (11), we have

$$\Pr\left(\tilde{\boldsymbol{v}}_{n+1} \leq y \middle| \tilde{x}_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right) = F_\upsilon\left(y\left|\tilde{\upsilon}_n, \tilde{r}_n, \delta_n, \tilde{\kappa}_j\right.\right), \tag{13}$$

which implies that

$$p_{\tilde{\boldsymbol{v}}_{n+1}}\left(\tilde{\upsilon}_{n+1}\left|\tilde{x}_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right.\right) = p_{\tilde{\kappa}_j, \upsilon}\left(\tilde{\upsilon}_{n+1}\left|\tilde{\upsilon}_n, \tilde{r}_n, \delta_n\right.\right). \tag{14}$$

Following similar steps, it is straightforward to show that

$$\begin{aligned}
&p_{\tilde{\boldsymbol{r}}_{n+1}}\left(\tilde{r}_{n+1}\left|\tilde{\upsilon}_{n+1}, \tilde{x}_{1:n}, \kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right.\right)\\
&\qquad = p_{\tilde{\kappa}_j, r}\left(\tilde{r}_{n+1}\left|\tilde{\upsilon}_{n+1}, \tilde{\upsilon}_n, \tilde{r}_n, \delta_n\right.\right).
\end{aligned} \tag{15}$$

Combining (10), (14) and (15), we have

$$\begin{aligned}
&p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right.\right)\\
&= \prod_{n=0}^{T-1} p_{\tilde{\kappa}_j, \upsilon}\left(\tilde{\upsilon}_{n+1}\left|\tilde{\upsilon}_n, \tilde{r}_n, \delta_n\right.\right) p_{\tilde{\kappa}_j, r}\left(\tilde{r}_{n+1}\left|\tilde{\upsilon}_{n+1}, \tilde{\upsilon}_n, \tilde{r}_n, \delta_n\right.\right)\\
&= p_{\tilde{\kappa}_j}\left(\tilde{x}_{1:T}; \delta_{1:T}\right).
\end{aligned} \tag{16}$$

The equation above implies that the Markov chain $\boldsymbol{\kappa} \rightarrow \left(\tilde{\boldsymbol{\kappa}}, \boldsymbol{\delta}_{1:T}\right) \rightarrow \tilde{\boldsymbol{x}}_{\boldsymbol{1:T}}$ holds. Thus, we have

$$\begin{aligned}
p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\tilde{\kappa}_j, \delta_{1:T}\right.\right) &= p_{\tilde{\boldsymbol{x}}_{1:T}}\left(\tilde{x}_{1:T}\left|\kappa_i, \tilde{\kappa}_j, \delta_{1:T}\right.\right)\\
&= p_{\tilde{\kappa}_j}\left(\tilde{x}_{1:T}; \delta_{1:T}\right).
\end{aligned}$$

Hence, the joint p.d.f. of the filter's output belongs to $\mathcal{M}$.

Appendix B
Proof of Theorem 2

Using the law of total expectation, $\mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_n\right\|^2\right]$ can be written as

$$\begin{aligned}
&\mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_n\right\|^2\right]\\
&\quad = \mathsf{E}_{\boldsymbol{\kappa}, \tilde{\boldsymbol{\kappa}}}\left[\mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_n\right\|^2 \middle| \boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right]\right]\\
&\quad = \sum_{\kappa_i, \tilde{\kappa}_j} \mathsf{E}\left[\left\|\boldsymbol{x}_n - \tilde{\boldsymbol{x}}_n\right\|^2 \middle| \boldsymbol{\kappa} = \kappa_i, \tilde{\boldsymbol{\kappa}} = \tilde{\kappa}_j\right]\\
&\qquad \times \pi\left(\tilde{\kappa}_j\left|\kappa_i\right.\right) \Pr\left(\boldsymbol{\kappa} = \kappa_i\right).
\end{aligned}$$

Thus, the objective function is linear in the randomization probabilities. The first and second constraints in (7) are linear and the privacy constraint is convex in the randomization probabilities [33]. These observations imply that the optimization problem (7) is convex.

Appendix C
Proof of Theorem 3

Using Fano's inequality [33], the error probability of any estimator of the true gains can be lower bounded as

$$\begin{aligned}
\Pr\left(\boldsymbol{\kappa} \neq \hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right)\right) &\geq \frac{\mathsf{H}\left[\boldsymbol{\kappa}\left|\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right.\right] - 1}{\log|\mathcal{K}|}.\\
&\overset{(a)}{=} \frac{\mathsf{H}\left[\boldsymbol{\kappa}\right] - \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right] - 1}{\log|\mathcal{K}|}.
\end{aligned} \tag{17}$$

where $\mathsf{H}\left[\boldsymbol{\kappa}\left|\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right.\right]$ is the discrete conditional entropy of $\boldsymbol{\kappa}$ given $\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}$ and (a) follows from the definition of mutual information. Next, we derive an upper bound on the mutual information between $\boldsymbol{\kappa}$ and $\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}$. In Appendix A, we show that the joint p.d.f. of $\tilde{\boldsymbol{x}}_{1:T}$ given $\boldsymbol{\delta}_{1:n}$, $\tilde{\boldsymbol{\kappa}}$ and $\boldsymbol{\kappa}$ only depends on $\tilde{\boldsymbol{\kappa}}$. Thus, the following Markov chain holds: $\boldsymbol{\kappa} \rightarrow \left(\tilde{\boldsymbol{\kappa}}, \boldsymbol{\delta}_{1:T}\right) \rightarrow \tilde{\boldsymbol{x}}_{1:T}$. Hence, we have

$$\begin{aligned}
\mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right] &\overset{(a)}{=} \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{x}}_{1:T}\right] + \mathsf{I}\left[\boldsymbol{\kappa}; \boldsymbol{\delta}_{1:T}\left|\tilde{\boldsymbol{x}}_{1:T}\right.\right]\\
&\overset{(b)}{=} \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{x}}_{1:T}\right]\\
&\overset{(c)}{\leq} \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{\kappa}}, \boldsymbol{\delta}_{\boldsymbol{1:T}}\right]\\
&\overset{(d)}{=} \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{\kappa}}\right] + \mathsf{I}\left[\boldsymbol{\kappa}; \boldsymbol{\delta}_{1:T}\left|\tilde{\boldsymbol{\kappa}}\right.\right]\\
&\overset{(e)}{=} \mathsf{I}\left[\boldsymbol{\kappa}; \tilde{\boldsymbol{\kappa}}\right]\\
&\leq I_0,
\end{aligned} \tag{18}$$

where (a) and (d) follow from the chain rule for mutual information [33], (b) and (e) follow from the fact that the driver steering command is deterministic. The inequality (c) follows from the data processing inequality [33] and the Markov chain $\boldsymbol{\kappa} \rightarrow \left(\tilde{\boldsymbol{\kappa}}, \boldsymbol{\delta}_{1:T}\right) \rightarrow \tilde{\boldsymbol{x}}_{1:T}$. The last inequality follows from the privacy constraint in optimization problem (7). Combining (17) and (18), we have

$$\Pr\left(\boldsymbol{\kappa} \neq \hat{\boldsymbol{\kappa}}\left(\tilde{\boldsymbol{x}}_{1:T}, \boldsymbol{\delta}_{1:T}\right)\right) \geq \frac{\mathsf{H}\left[\boldsymbol{\kappa}\right] - I_0 - 1}{\log|\mathcal{K}|}.$$

References

[1] K. Koscher *et al.*, "Experimental security analysis of a modern automobile," in *Proc. IEEE Symp. Secur. Privacy*, May 2010, pp. 447–462.

[2] S. Checkoway *et al.*, "Comprehensive experimental analyses of automotive attack surfaces," in *Proc. 20th USENIX Conf. Secur.*, 2011, p. 6.

[3] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, Jan. 2015.

[4] C. Schellenberger and P. Zhang, "Detection of covert attacks on cyber-physical systems by extending the system dynamics with an auxiliary system," in *Proc. IEEE 56th Annu. Conf. Decis. Control (CDC)*, Dec. 2017, pp. 1374–1379.

[5] G. Park, H. Shim, C. Lee, Y. Eun, and K. H. Johansson, "When adversary encounters uncertain cyber-physical systems: Robust zero-dynamics attack with disclosure resources," in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, Dec. 2016, pp. 5090–15085.

[6] A. Hoehn and P. Zhang, "Detection of covert attacks and zero dynamics attacks in cyber-physical systems," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2016, pp. 302–307.

[7] X. He, W. P. Tay, and M. Sun, "Privacy-aware decentralized detection using linear precoding," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Jul. 2016, pp. 1–5.

[8] M. Sun and W. P. Tay, "Privacy-preserving nonparametric decentralized detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2016, pp. 6270–6274.

[9] X. He and W. P. Tay, "Multilayer sensor network for information privacy," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2017, pp. 6005–6009.

[10] J. Liao, L. Sankar, V. Y. F. Tan, and F. P. Calmon, "Hypothesis testing in the high privacy limit," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2016, pp. 649–656.

[11] Z. Li and T. J. Oechtering, "Privacy-aware distributed Bayesian detection," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1345–1357, Oct. 2015.

[12] Z. Li and T. J. Oechtering, "Privacy-constrained parallel distributed Neyman-Pearson test," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 1, pp. 77–90, Mar. 2017.

[13] K. Kalantari, L. Sankar, and O. Kosut, "On information-theoretic privacy with general distortion cost functions," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2865–2869.

[14] Y. O. Basciftci, Y. Wang, and P. Ishwar, "On privacy-utility tradeoffs for constrained data release mechanisms," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Jan. 2016, pp. 1–6.

[15] F. du P. Calmon and N. Fawaz, "Privacy against statistical inference," in *Proc. IEEE 50th Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, Oct. 2012, pp. 1401–1408.

[16] B. Moraffah and L. Sankar, "Information-theoretic private interactive mechanism," in *Proc. 53rd Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, Sep. 2015, pp. 911–918.

[17] E. Nekouei, H. Sandberg, M. Skoglund, and K. H. Johansson, "Optimal privacy-aware estimation," *IEEE Trans. Autom. Control*, early access, May 6, 2021, doi: 10.1109/TAC.2021.3077868.

[18] R. Mochaourab and T. J. Oechtering, "Private filtering for hidden Markov models," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 888–892, Jun. 2018.

[19] P. Venkitasubramaniam, "Privacy in stochastic control: A Markov decision process perspective," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 381–388.

[20] E. Nekouei, T. Tanaka, M. Skoglund, and K. H. Johansson, "Information-theoretic approaches to privacy in estimation and control," *Annu. Rev. Control*, vol. 47, pp. 412–422, Jan. 2019.

[21] J. L. Ny and G. J. Pappas, "Differentially private filtering," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 341–354, Feb. 2013.

[22] H. Sandberg, G. Dán, and R. Thobaben, "Differentially private state estimation in distribution networks with smart meters," in *Proc. 54th IEEE Conf. Decis. Control (CDC)*, Dec. 2015, pp. 4492–4498.

[23] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design," *Automatica*, vol. 81, pp. 221–231, 2017.

[24] Y. Mo and R. M. Murray, "Privacy preserving average consensus," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 753–765, Feb. 2017.

[25] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud, "Differential privacy in linear distributed control systems: Entropy minimizing mechanisms and performance tradeoffs," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 1, pp. 118–130, Mar. 2017.

[26] E. Nekouei, H. Sandberg, M. Skoglund, and K. H. Johansson, "A model randomization approach to statistical parameter privacy," 2021, *arXiv:2105.10664*.

[27] R. Rajamani, *Vehicle Dynamics and Control*. New York, NY, USA: Springer, 2011.

[28] M. Pirani *et al.*, "Resilient corner-based vehicle velocity estimation," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 2, pp. 452–462, Mar. 2018.

[29] X. Huang, H. Zhang, G. Zhang, and J. Wang, "Robust weighted gain-scheduling H_∞ vehicle lateral motion control with considerations of steering system backlash-type hysteresis," *IEEE Trans. Control Syst. Technol.*, vol. 22, pp. 1740–1753, 2014.

[30] H. Zhang and J. Wang, "Vehicle lateral dynamics control through AFS/DYC and robust gain-scheduling approach," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 489–494, Jan. 2016.

[31] R. S. C. Campolo and A. Molinaro, *Vehicular ad hoc Networks Standards, Solutions, and Research*. Boston, MA, USA: Springer, 2015.

[32] G. Dimitrakopoulos, *Current Technologies in Vehicular Communication*. New York, NY, USA: Springer, 2017.

[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2006.

[34] C. Dwork and A. Roth, *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends in Theoretical Computer Science, 2014.