A Security Mechanism Against Inference Attacks on Networked Systems

Ehsan Nekouei, Mohammad Pirani, Chuanghong Weng, and Michaël Antonie van Wyk

***Abstract*—This paper develops a security mechanism against inference attacks for industrial systems where an adversary with access to the states of a (linear or nonlinear) system attempts to infer the system model using the state values. Under an inference attack, an adversary with access to the sensor measurements of the system attempts to infer the system's parameters. The proposed security mechanism consists of two components: (*i*) a collection of feedback control gains, (*ii*) a randomized gain selection policy. To mitigate the inference attack, the gain selection policy randomly selects a feedback gain from the set of available feedback control gains at regular intervals. We cast the optimal design of the gain selection policy as an optimization problem such that (*i*) quadratic control cost is minimized and (*ii*) the uncertainty level of the adversary about selected control gain is maximized. In our formulation, the uncertainty level of the adversary about the control gain is captured by the Kullback-Leibler (KL) divergence between a uniform distribution and the posterior distribution of the feedback gains, given the history of the system states. We first derive the backward Bellman optimality equation for the gain selection problem and study the structural properties of the optimal gain selection policy. Our results show that the optimal gain selection policy only depends on the current state of the system, rather than the entire history of the states, which renders the optimal gain selection problem to a nonlinear Markov decision process. Next, we derive a policy gradient theorem for the gain selection problem, which provides an expression for the gradient of the objective function of the gain selection problem with respect to the parameter of a stationary (time-invariant) policy. The policy gradient theorem allows us to develop a stochastic gradient descent algorithm for computing an optimal policy. We finally demonstrate the effectiveness of our results for different linear and nonlinear systems. Our results indicate that the proposed security mechanism significantly decreases the inference ability of the adversary, while having a negligible impact on the control cost.**

Received 16 August 2024; revised 24 March 2025, 26 June 2025, and 24 September 2025; accepted 28 September 2025. Date of publication 6 October 2025; date of current version 9 October 2025. This work was supported in part by Hong Kong Research Grants Council under Project CityU 21208921 and Project CityU 11210024 and in part by Chow Sang Sang Group Research Fund. The associate editor coordinating the review of this article and approving it for publication was Dr. Sharif Abuadbba. *(Corresponding author: Ehsan Nekouei.)*

Ehsan Nekouei and Chuanghong Weng are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong (e-mail: enekouei@cityu.edu.hk; cweng7-c@my.cityu.edu.hk).

Mohammad Pirani is with the Department of Mechanical Engineering, University of Ottawa, Ottawa, ON K1H 8M5, Canada (e-mail: mpirani@uottawa.ca).

Michaël Antonie van Wyk is with the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg 2000, South Africa, and also with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, SAR, China (e-mail: Anton.vanWyk@wits.ac.za).

Digital Object Identifier 10.1109/TIFS.2025.3616608

***Index Terms*—Cyber-physical security, inference attack, Kullback-Leibler (KL) divergence, moving target framework.**

I. Introduction

A. Motivation

THE role of humans is integrated into industrial systems in numerous applications. Humans can engage with an industrial system either as users or operators. In certain situations, however, these human roles can shift to that of an adversary, aiming to compromise the system, gather information, or cause a specific impact. To counteract such threats, three primary defense mechanisms are typically employed: prevention, detection, and mitigation. Among these, prevention takes precedence as the first line of defense to safeguard the system. One primary approach to preventing attacks in industrial systems is to enhance the ambiguity of system states and parameters. However, increasing ambiguity can compromise the system's performance and optimality. In this paper, we explore a meticulously designed randomization of control gains. This approach aims to maximize the adversary's uncertainty without compromising the system's performance.

Inference attacks are an important class of cyber-physical attacks, in which an adversary with access to the sensor measurements of a system attempts to infer the parameters of the system. The knowledge of the system parameters allows an attacker to launch model-based stealthy attacks on control systems. For instance, an attacker with the knowledge of system parameters can perform a zero dynamic attack on a control system in which the attack signal is hidden in the output-nulling space of the system.

B. Related Work

As industrial systems became more integrated with communication platforms, concerns about their security were raised. Security methods in these systems can be classified into three levels: prevention algorithms, detection algorithms, and mitigation algorithms, see [1] and references therein. The first layer of defense is to prevent the attack from happening. In many cases, however, it is not possible to prevent all attacks (e.g., if attackers exploit subtle "zero-day" vulnerabilities [2], or rely on insider threats [3]). In those cases, the second layer comes into play, which aims to detect and isolate the attack [4], [5]. Besides attack detection, the target system must be resilient enough to withstand the attacks or mitigate the impact of the attack [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17].

Our focus in this paper is on a *prevention mechanism* and therefore belongs to the first defense layer. When estimation,

control, or actuation tasks in an industrial system are performed by an untrusted party, the risk of private information leakage becomes significant. To address this, a variety of prevention-oriented strategies have been developed. Algebraic methods in network coding have shown how communication can be made both efficient and resilient to failures, ensuring that data continues to flow reliably even under disruptions [18]. Cryptographic approaches, such as semi-homomorphic encryption, allow controllers to process encrypted sensor data without ever exposing the underlying information, thus preserving confidentiality while maintaining system stability [19]. Differential privacy provides another rigorous framework, guaranteeing that released data, whether in statistical analyses or dynamic filtering of signals, offers strong privacy protection against adversaries with side information [20], [21].

Randomization techniques add further unpredictability, making it computationally difficult for attackers to infer sensitive system properties [22]. Complementary to these approaches, moving target defenses deliberately vary system dynamics over time, frustrating attempts to build accurate models for stealthy attacks [23]. Finally, information-theoretic formulations provide fundamental limits on what an adversary can infer, enabling privacy-aware decision making in applications such as smart grids and smart buildings [24]. Together, these mechanisms illustrate the breadth of available tools for preventing attacks and form the basis of the first layer of defense in modern networked control systems.

One practical prevention method is altering critical system parameters periodically to prevent attackers from inferring them. A recent approach, called moving target defense, achieves this by introducing random, time-varying parameters into a system. This limits the attacker's knowledge of the model, making it difficult to devise covert attack strategies, while the dynamic nature of the system further discourages adaptive adversaries. The same principle has been applied in learning frameworks to improve security and privacy. For instance, in decentralized federated learning, moving target defense techniques such as random neighbor selection and IP/port switching have been combined with encryption to protect against communication-based threats like eavesdropping and eclipse attacks, demonstrating both strong robustness and practical efficiency [25]. In deep learning, feature randomization has been used to resist adversarial attack transferability across models by altering the feature space during training and testing. This prevents attackers from crafting adversarial inputs that generalize across models, significantly strengthening resilience against a wide range of attack methods [26]. Together, these studies highlight the versatility of moving target defense in securing not only cyber-physical systems but also modern machine learning platforms.

One key parameter within the system that can be directly adjusted by the designer is the set of control gains, making them suitable for randomization to obscure the true values. Gain randomization maximizes system ambiguity by introducing pseudo-gains, limiting what an attacker can infer. In [27], a general model randomization framework was proposed using randomizers and nonlinear transformations to enforce parameter privacy under information-theoretic constraints. Extending this idea to control, [28] developed a randomized filtering strategy for active steering systems, showing that pseudo-gains and nonlinear mappings can effectively conceal true controller parameters while outperforming additive noise mechanisms. Similar concepts have been applied in other domains: verifiable randomization has been used to strengthen local differential privacy protocols against malicious manipulation [29], while randomized masking has been shown to improve robustness of text classifiers against adversarial perturbations [30]. These results collectively highlight the versatility of parameter randomization and motivate its study for closed-loop control through optimal control and information-theoretic frameworks.

C. Contributions

In this paper, we propose a security mechanism against inference attacks on industrial systems. In our setup, an adversary with access to the history of the state measurements of a system attempts to infer the system's model. As a security countermeasure against such an inference attack, we develop a gain randomization framework wherein a gain selection policy randomly at each time step selects a feedback gain from a collection of predefined feedback gains. We formulate the gain selection policy as an optimal control problem where the objective is to minimize a quadratic control cost and maximize the uncertainty level of the adversary about the system model, captured by the Kullback-Leibler divergence between a uniform distribution and the posterior distribution of the feedback gains given the state measurements. We then study the structural properties of the optimal gain selection policy and develop a numerical algorithm for executing an optimal policy. Our main contributions can be summarized as follows:

1) We derive Bellman's optimality principle for the optimal gain selection policy. Here, our results indicate that the optimal gain selection policy is only a function of the current state of the system rather than the entire history of the state measurement. As a result, the gain selection policy becomes a nonlinear Markov decision process. We also show that the optimal gain selection policy can be computed backward in time by solving a series of convex optimization problems.
2) We derive a policy gradient theorem that provides an expression for the gradient of the objective function of the gain randomization problem with respect to the parameters of a stationary policy.
3) Using the policy gradient theorem, we develop a stochastic gradient descent algorithm for computing an optimal gain selection policy.
4) The computation of the gradient depends on certain likelihood functions, e.g., see equation (8), which may not admit a closed-form when the system is nonlinear. To solve this problem, we show that the term, which involves the likelihood functions, can be indirectly computed by solving an optimization problem.

The rest of this paper is organized as follows. The next section presents our system model and the optimal gain selection problem for linear systems. Sections III is dedicated to the structural properties of the optimal gain selection policy and its numerical computation in the case of linear systems. There results are extended to nonlinear systems in Section IV.

Our numerical results for both linear and nonlinear systems are presented in Section V, followed by the concluding remarks in Section VI.

II. System Model and Problem Formulation

Consider a linear networked system in which the system state evolves according to

$$X_{k+1} = AX_k + BU_k + W_k, \tag{1}$$

where X_k and W_k are the system's state and the process noise at time-step k, respectively, the matrices A, B are the system parameters, and U_k is the control input at time k. The process noise is modeled as a sequence of independent and identically distributed random vectors. The initial state of the system, *i.e.*, X_1 is assumed to be independent of $\{W_k\}_k$.

In most industrial systems, sensors and controllers are not collocated. Thus, the sensor measurements are sent to the control unit via a communication network. For example, at time k, the sensor measurement X_k is sent from the sensor to the controller. The control unit then selects the control input at time k, *i.e.,* U_k, based on the state measurements X_k. The control input U_k can be equal to KX_k where K is a matrix of appropriate dimension. In this case, the control input is a linear combination of the states of the system where the entries of K determine the gains (weights) in the linear combination. Thus, in what follows, we refer to K as the control gain. However, the control input can be a nonlinear function of the system state in general.

A. Attack Model

We consider an adversary that eavesdrops on the communication between the sensor and control unit. Thus, the adversary has access to the entire history of the system states $\{X_1, \ldots, X_T\}$. Furthermore, the adversary is interested in the knowledge of the system model, *i.e.,* the system parameters. Since the parameters A and B in (1) depend on the physical properties of the system, we assume that they are known by the adversary. Under a time-invariant control law, *e.g.,* $U_k = KX_k$, the adversary can infer the feedback gain K based on the system states $\{X_1, \ldots, X_T\}$, *e.g.,* using the least squares estimator. This gives the adversary access to the complete system model, *i.e.,* (A, B, K), which allows the adversary to launch powerful model-based attacks on the system. In the next subsection, we propose a security mechanism against the inference attack.

B. Gain Randomization Mechanism

To counteract the inference attack, we consider a security mechanism in which the feedback gain of the system is regularly changed. More precisely, the proposed security mechanism against inference attacks consists of two components: (*i*) a predefined set of gains denoted by $\{K_1, \ldots, K_M\}$, (*ii*) a randomized gain selection policy which randomly selects a feedback gain from the set $\{K_1, \ldots, K_M\}$ at every time step. More specifically, the control gain at time k is selected based on the probability distribution $\pi_k(\cdot|I_k)$ where

$$I_k = \{x_1, \ldots, x_k\}, \tag{3}$$

is the entire history of the system states up to time k. Thus, under $\pi_k(\cdot|I_k)$, the gain K_i is selected with probability $\pi_k(i|I_k)$. In the rest of the paper, we refer to $\pi = \{\pi_k(\cdot|I_k)\}_k$ as the *gain selection policy*.

Fig. 1 on the next page provides a high-level description of this mechanism. In this figure, L_k denotes a random variable that encodes the index of the selected control gain at time k, *i.e.*, $L_k = i$ if the control gain K_i is selected at time k. Thus, the control input of the system at time k can be written as $U_k = K_{L_k} X_k$.

In this paper, we assumed the adversary uses a maximum likelihood estimator to infer the feedback gains. To measure the inference ability of the adversary at time k, we use the posterior distribution of L_k given $X_{1:k+1} = x_{1:k+1}$, *i.e.*, $\mathsf{Pr}\left(L_k | x_{1:k+1}\right)$, where we used the shorthand notations $X_{1:k+1} = \{X_1, \ldots, X_{k+1}\}$ and $x_{1:k+1} = \{x_1, \ldots, x_{k+1}\}$. Note that the adversary has access to the entire history of the system states $x_{1:k+1} = \{x_1, \ldots, x_{k+1}\}$, and can form the posterior distribution $\mathsf{Pr}\left(L_k | x_{1:k+1}\right)$ using $x_{1:k+1}$. Note that $\mathsf{Pr}\left(L_k | x_{1:k+1}\right)$ represents the belief of the adversary about the control gain used at time k. Moreover, the adversary has maximum ambiguity about the feedback gain of the system when the conditional distribution of L_k given $X_{1:k+1} = x_{1:k+1}$ is uniformly distributed. The inference attack can be formally expressed as

$$\arg \max_{i=1,\ldots,M} \mathsf{Pr}\left(L_k = i | x_{1:k+1}\right),$$

where the adversary first forms the posterior probability distribution of the selected gain given the available measurements and then selects the gain associated with the highest posterior probability.

To capture the impact of the gain selection policy on the inference ability of the adversary at time k, we use the Kullback-Leibler (KL) divergence between a uniform distribution and the posterior distribution of L_k given $X_{1:k+1} = x_{1:k+1}$, which is defined as

$$\mathsf{D}\left[\mathsf{Uni}_M(j) \,\middle\|\, \mathsf{Pr}\left(L_k = j | x_{1:k+1}\right)\right] = \frac{1}{M} \sum_j \log \frac{\frac{1}{M}}{\mathsf{Pr}\left(L_k = j | x_{1:k+1}\right)} \tag{4}$$

where $j \in \{1, \ldots, M\}$ and $\mathsf{Uni}_M(\cdot)$ is a uniform probability mass function on $\{1, \ldots, M\}$. Note that the above KL divergence is always non-negative and is zero when the conditional distribution of L_k given $X_{1:k+1} = x_{1:k+1}$ is uniformly distributed, which corresponds to the maximum ambiguity level of the adversary about the feedback gain. Thus, it is desirable to design the gain selection policy such that the KL divergence above is as small as possible by proper choice of the gain

$$\min_{\{\pi_k(\cdot|I_k)\}_k} \quad \sum_{k=1}^{T} \mathsf{E}\left[X_k^\top Q X_k + U_k^\top R U_k + \lambda \mathsf{D}\left[\mathsf{Uni}_M(j) \,\middle\|\, \mathsf{Pr}\left(L_k = j | X_{1:k+1}\right)\right]\right] + \mathsf{E}\left[X_{T+1}^\top Q X_{T+1}\right] \tag{2}$$

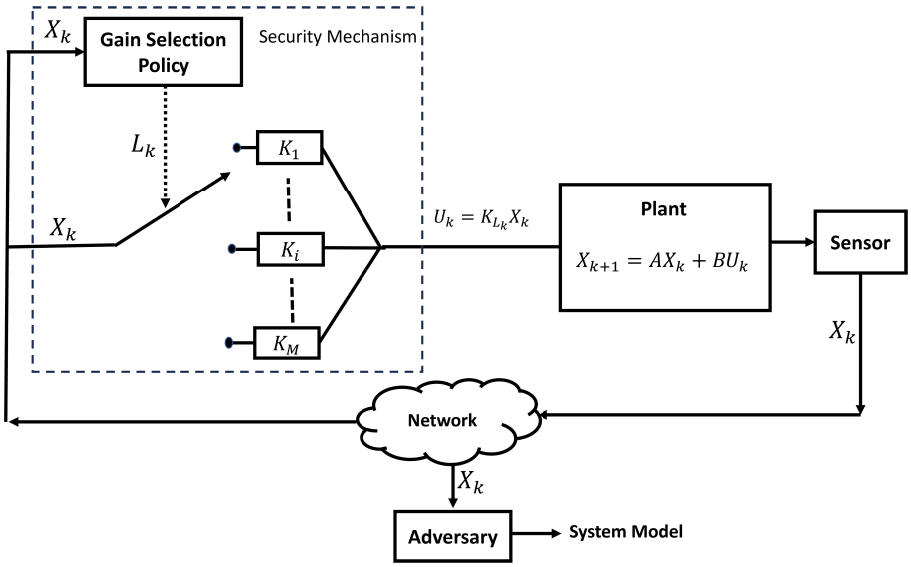


Fig. 1. The schematic representation of the proposed security mechanism against inference attack on a linear networked system.

selection policy. It is straightforward to show that for the linear system in (1) $\mathsf{Pr}\left(L_k = j | x_{1:k+1}\right)$ is always nonzero if the probability density function of W_k is nonzero everywhere.

Remark 1: It is well-known that the KL divergence is not symmetric in its arguments, *i.e.,* $\mathsf{D}\left[\mathsf{P} \| \mathsf{Q}\right]$ is not necessarily equal to $\mathsf{D}\left[\mathsf{Q} \| \mathsf{P}\right]$ for two probability distributions P and Q. The choice of direction for the KL divergence in (4) significantly simplifies the numerical calculation of the optimal policy. However, in the alternative direction, the KL divergence will be a highly nonlinear function of the policy which may significantly complicate the computations.

C. Optimal Design of the Gain Selection Policy

The gain selection policy $\pi = \{\pi_k(\cdot | I_k)\}_k$ is designed to achieve two objectives (*i*) to minimize a quadratic control cost, (*ii*) to maximize the uncertainty of the adversary about the control gain of the system. To this end, we cast the optimal design of the gain section policy as the optimization problem as in (2), shown at the bottom of the previous page, where R and Q are positive semi-definite matrices and λ is a positive constant. The term $X_k^\top Q X_k + U_k^\top R U_k$ in (2) captures the impact of the gain selection policy on the states and control input of the system, whereas the term $\mathsf{D}\left[\mathsf{Uni}_M(j) \| \mathsf{Pr}\left(L_k = j | X_{1:k+1}\right)\right]$ captures the impact of the gain selection policy on the uncertainty level of the adversary about the feedback gain. Note that the adversary has maximum uncertainty about L_k when $\mathsf{Pr}\left(L_k | X_{1:k+1}\right)$ is uniformly distributed. Therefore, the optimization problem (2) finds the optimal gain selection policy such that the control cost is minimized while ensuring the adversary cannot reliably infer the control gain.

Remark 2: The choice of the uniform distribution in (4) is motivated by the fact that the uniform distribution results in maximum entropy, *i.e.,* maximum ambiguity for the adversary. In practice, the uniform distribution in the KL divergence in (4) can be replaced with a reference probability measure π_k^{ref}. In this case, the objective of the gain selection problem is to ensure that the posterior distribution $\mathsf{Pr}\left(L_k | x_{1:k+1}\right)$ is close to the reference distribution π_k^{ref}, which can be selected by the system designer based on the physical properties of the system as well as the performance and security requirements.

III. Structure and Computation of the Optimal Gain Selection Policy

In this subsection, we first derive the optimality equations for the gain selection problem, which is used to study the structural properties of the optimal gain selection policy. We then derive a numerical algorithm for computing an optimal policy.

Theorem 1: Let $\pi_k^\star(\cdot | \cdot)$ denote the optimal gain selection policy at time k. Then, the following statements are true.

1) Without loss of optimality, the search space of the optimization problem (2) can be reduced to the collection of the policies of the form $\{\pi_k(\cdot | x_k)\}_{k=1}^T$, in which the gain selection policy at time k only depends on X_k. That is, X_k is the sufficient statistics for decision-making in the gain selection problem.
2) The optimal gain selection policy $\pi_k^\star(\cdot | \cdot)$ is only a function of X_k.
3) The optimal gain selection policy satisfies the backward optimality as in (5), shown at the bottom of the next page, where $V_k^\star(\cdot)$ is the optimal value function at time k, the terminal value function is given by $V_{T+1}^\star\left(x_{T+1}\right) = x_{T+1}^\top Q x_{T+1}$, and $p\left(x_{k+1} | x_k, L_k = j\right)$ is the conditional density of X_{k+1} given $X_k = x_k, L_k = j$.
4) The optimization problem in (5) is convex.
5) The optimal gain selection problem in (2) is a nonlinear Markov decision process (MDP).

Proof: See Appendix A. ∎

According to Theorem 1, the optimal gain selection policy satisfies the backward Bellman optimality principle in (5), which can be used to (*i*) derive value-based methods for computing an optimal policy, (*ii*) recursively compute the optimal value function associated with the optimal policy. A main consequence of Theorem 1 is that the optimal gain selection problem is a feedback control problem in the form of an MDP. This is due to the fact that the choice of the

feedback gain at time k affects the evolution of the system state, and consequently, the next state (X_{k+1}) will depend on the gain selection policy at time k. Thus, the impact of the gain selection policy at time k $(\pi_k(\cdot|\cdot))$ on the future states must be taken into account in the design of optimal gain selection policy.

Different from standard MDPs, the optimal gain selection problem (2) is a nonlinear MDP, where the objective function is a nonlinear function of the policy due to the KL divergence term in the objective function. As a result, the optimal gain selection policy will be a randomized function of the current state of the system. However, in a standard MDP, the objective function will be linear in the policy, and as a result, the optimal policy will be a deterministic function of the state [31].

The feasible set of the optimal gain selection problem (2) is the set of all conditional probabilities of the form $\{\pi_k(\cdot|x_1,\ldots,x_k)\}_{k=1}^T$ which depend on the history of the states. However, based on Theorem 1, it is optimal to reduce the search space for the optimal policy to the policies which only depend on the current state, *i.e.*, $\{\pi_k(\cdot|x_k)\}_{k=1}^T$, which is substantially smaller than the space of the policies of the form $\{\pi_k(\cdot|x_1,\ldots,x_k)\}_{k=1}^T$. This observation implies that X_k is the sufficient statistic for decision-making in the optimal gain selection problem. This observation also significantly reduces the computational complexity of the search for the optimal policy as it allows the design of stationary (time-invariant) optimal policies, which is impossible when the optimal policy depends on the entire history of state trajectory. Finally, the optimal value function at time k only depends on x_k rather than $x_1,\ldots,x_k$, which is especially helpful when the value-based methods are used to compute an optimal policy.

A. Policy Gradient Theorem

In this subsection, we develop a policy gradient algorithm for finding the optimal gain selection policy. To this end, let $\pi_\theta(\cdot|\cdot)$ denote a policy parameterized by θ. Thus, at time k, a feedback gain is randomly selected according to the conditional probability distribution $\pi_\theta(\cdot|x_k)$, which depends on the state at time k. In the next theorem, we derive an expression for the gradient of the objective function of the optimization problem (2) with respect to θ. Using this expression, we then develop a policy gradient algorithm for solving (2).

Theorem 2: Consider a gain randomization policy of the form $\pi_\theta(\cdot|\cdot)$ which is parameterized by $\theta\in\mathbb{R}^p$. Let J_θ denote the value of the objective function of the optimization problem (2) under $\pi_\theta(\cdot|\cdot)$. Then, the gradient of J_θ with respect to θ can be written as in (6), shown at the bottom of the page, where $\Phi_\theta(x_{1:T+1})$ is given by

$$\Phi_\theta(x_{1:T+1}) = \frac{1}{M}\sum_{k,j}\log\frac{\sum_i p(x_{k+1}|x_k,L_k=i)\pi_\theta(i|x_k)}{p(x_{k+1}|x_k,L_k=j)\pi_\theta(j|x_k)}. \tag{8}$$

Proof: See Appendix B. ∎

According to Theorem 2, the gradient of the objective function with respect to θ consists of two terms. The first term in (6) follows from the standard Policy Gradient Theorem of MDPs, where as the second term in (6) is due to the nonlinear term introduced in the objective function by the KL divergence. Note that the second term vanishes when λ is equal to zero and the optimization problem (2) reduces to a standard MDP.

B. Stochastic Gradient Descent Algorithm

In this subsection, we use Theorem 2 to develop a stochastic gradient algorithm for computing an optimal policy. Let s denote the iteration index of the algorithm. Also, let $\pi_{\theta_s}(\cdot|\cdot)$ and θ_s denote the gain selection policy and the policy parameter at iteration s, respectively. To update the policy parameter at iteration s, we first use the current policy $\pi_{\theta_s}(\cdot|\cdot)$ to generate N samples of the system state trajectory, and their corresponding feedback gain indices by sampling from the policy $\pi_{\theta_s}(\cdot|\cdot)$.

$$V_k^\star(x_k) = \min_{\pi_k(\cdot|x_k)} x_k^\top Q x_k + \sum_{j=1}^M x_k^\top K_j^\top R K_j x_k \pi_k(j|x_k) + \log\frac{1}{M} + \frac{1}{M}\sum_{j=1,l=1}^M \mathsf{E}\left[V_{k+1}^\star(X_{k+1}) - \lambda\log\frac{p(X_{k+1}|x_k,L_k=j)\pi_k(j|x_k)}{\sum_i p(X_{k+1}|x_k,L_k=i)\pi_k(i|x_k)}\,\middle|\, X_k=x_k, L_k=l\right]\pi_k(l|x_k) \tag{5}$$

$$\nabla_\theta J_\theta = \mathsf{E}\left[\left(X_{T+1}^\top Q X_{T+1} + \sum_{k=1}^T X_k^\top Q X_k + U_k^\top R U_k\right)\sum_{k=1}^T \nabla_\theta\log\pi_\theta(L_k|X_k)\right] + \mathsf{E}\left[\Phi_\theta(X_{1:T+1})\sum_{k=1}^T\nabla_\theta\log\pi_\theta(L_k|X_k) - \frac{1}{M}\sum_{k=1}^T\sum_{j=1}^M\nabla_\theta\log\pi_\theta(j|X_k)\right] \tag{6}$$

$$\theta_{s+1} = \theta_s - \mu_s\frac{1}{N}\sum_{i=1}^N\left(x_{T+1}^{s,i\,\top} Q x_{T+1}^{s,i} + \sum_{k=1}^T x_k^{s,i\top} Q x_k^{s,i} + u_k^{s,i\top} R u_k^{s,i} + \Phi_{\theta_s}(x_{1:T+1}^{s,i})\right)\sum_{k=1}^T\nabla_\theta\log\pi_\theta(l_k^{s,i}|x_k^{s,i})\Big|_{\theta=\theta_s} + \mu_s\frac{1}{N}\sum_{i=1}^N\frac{1}{M}\sum_{k=1}^T\sum_{j=1}^M\nabla_\theta\log\pi_\theta(j|x_k^{s,i})\Big|_{\theta=\theta_s} \tag{7}$$

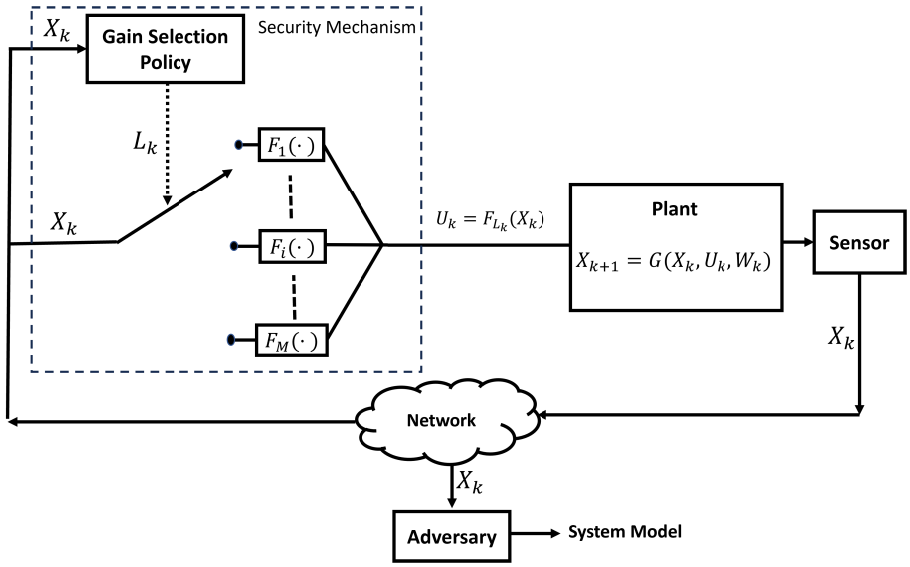


Fig. 2. The schematic representation of the proposed security mechanism for a nonlinear networked system.

Algorithm 1 Policy Gradient Algorithm

Initialize iteration index $s = 0$ and policy parameter θ_0.
repeat
  for $i = 1$ to N **do**
    Generate $x_1^{s,i}$
    for $k = 1$ to T **do**
      Generate $l_k^{s,i}$ by sampling from $\pi_{\theta_s}\left(\cdot \mid x_k^{s,i}\right)$
      $x_{k+1}^{s,i} \leftarrow A x_k^{s,i} + B K_{l_k^{s,i}} + w_k^{s,i}$
    end for
  end for
  Compute θ_{s+1} using (7)
  $s \leftarrow s + 1$
until θ_s converges

Let $x_{1:T+1}^{s,i} = \left\{x_1^{s,i}, \ldots, x_{T+1}^{s,i}\right\}$ and $l_{1:T}^{s,i} = \left\{l_1^{s,i}, \ldots, l_T^{s,i}\right\}$ denote the ith sample trajectory of the system state at iteration s and its corresponding trajectory of feedback gain indices, respectively. We then update the policy parameter according to as in (7), shown at the bottom of the previous page, where μ_s is the step-size at time-step s. Note that the expectation in (6) is costly to compute numerically. Therefore, in (7), we approximate the expectation using N sample paths of system state and the selected gains. Different steps of the policy gradient algorithm for solving (2) are shown in Algorithm 1.

$$\sup_{f(\cdot,\cdot,\cdot)\in L_+^3} \mathsf{E}\left[\log f\left(\tilde{L}, X_k, X_{k+1}\right)\right] - \mathsf{E}\left[f\left(L_k, X_k, X_{k+1}\right)\right] + 1 \quad (9)$$

IV. Extension to Nonlinear Systems

In this section, we extend the developed gain randomization framework to stochastic nonlinear systems of the form

$$X_{k+1} = G\left(X_k, U_k, W_k\right), \quad (10)$$

where $G(\cdot,\cdot,\cdot)$ is a nonlinear mapping, W_k is the process noise, and U_k is the control input of the system at time k, which can be in the form of a state feedback controller $U_k = F(X_k)$. We assume that the adversary has access to the entire history of the system states $\left\{x_1, \ldots, x_{T+1}\right\}$ as well as the nonlinear mapping $G(\cdot,\cdot,\cdot)$.

Let $\{F_1(x) \ldots, F_M(x)\}$ denote a collection of state feedback controllers, where $F_i(x)$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. To counteract the inference attack, we consider a security mechanism in which the controller is randomly changed at different time steps, as shown in Fig. 2. More specifically, the feedback control law at time k is selected based on the probability distribution $\pi_k(\cdot \mid I_k)$ where $I_k = \{x_1, \ldots, x_k\}$.

Remark 3: The optimal gain selection policy for nonlinear systems can be readily obtained by extending the Bellman optimality condition in (5) to the nonlinear case by substituting the term $x_k^\top K_j^\top R K_j x_k \pi_k(j \mid x_k)$ in (5) with $F_j(x_k)^\top R F_j(x_k)\, \pi_k(j \mid x_k)$.

A major challenge in the nonlinear case is the numerical computation of the gradient of the objective function with respect to θ in (7) for time-invariant policies of the form $\pi_{\theta_s}(\cdot \mid \cdot)$. This is because the gradient expression in (7) depends on the function $\Phi_\theta\left(x_{1:T+1}\right)$ in (7) which is defined as

$$\Phi_\theta\left(x_{1:T+1}\right) = \frac{1}{M} \sum_{k,j} \log \frac{\sum_i p\left(x_{k+1} \mid x_k, L_k = i\right) \pi_\theta(i \mid x_k)}{p\left(x_{k+1} \mid x_k, L_k = j\right) \pi_\theta(j \mid x_k)}.$$

Note that $\Phi_\theta\left(x_{1:T+1}\right)$ depends on the conditional distribution $p\left(x_{k+1} \mid x_k, L_k = i\right)$, which in the linear case, can be easily obtained by changing the mean of the distribution of W_k to $(A + BK_i)\, x_k$. However, the distribution $p\left(x_{k+1} \mid x_k, L_k = i\right)$ may not have a closed-form expression in the nonlinear as it depends on the feedback law $F_l(x_k)$, the nonlinear mapping $G(\cdot,\cdot,\cdot)$ and the distribution of W_k.

To tackle this problem, we first express the function $\Phi_\theta\left(x_{1:T+1}\right)$ as

$$\Phi_\theta\left(x_{1:T+1}\right) = \sum_k \eta_\theta\left(x_{k+1}, x_k\right), \quad (11)$$

where $\eta_\theta\left(x_{k+1}, x_k\right)$ is given by

$$\eta_\theta\left(x_{k+1}, x_k\right) = \frac{1}{M}\sum_j \log \frac{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_\theta(i|x_k)}{p\left(x_{k+1}|x_k, L_k = j\right)\pi_\theta(j|x_k)}.$$

The next Theorem shows that $\eta_\theta(\cdot,\cdot)$ can be indirectly computed by solving an optimization problem.

Theorem 3: Let $f^\star(\cdot,\cdot,\cdot)$ denote the optimal solution of the optimization problem (9) where L_+^3 is the space of measurable, three-dimensional positive functions, and $\tilde{L}$ is a uniformly distributed random variable over $\{1,\ldots,M\}$ which is independent of L_k, X_k, X_{k+1}. Then, we have

$$\eta_\theta\left(x_{k+1}, x_k\right) = \log(M) + \frac{1}{M}\sum_j \log f^\star\left(j, x_k, x_{k+1}\right).$$

Based on Theorem 3, the function $\eta_\theta(\cdot,\cdot)$ can be indirectly computed by solving the optimization problem (9). Note that the optimization problem (9) can be efficiently solved using stochastic gradient method. Thus, in the nonlinear case, we can first compute $\eta_\theta(\cdot,\cdot)$ by solving (9). Using $\eta_\theta(\cdot,\cdot)$, we can then evaluate the function $\Phi_\theta\left(x_{1:T+1}\right)$ and the gradient of the objective function with respect to the policy parameter θ. Different steps of the policy gradient algorithm for the nonlinear systems are summarized in Algorithm 2.

Algorithm 2 Policy Gradient Algorithm for NonlinearSystems

Initialize iteration index $s = 0$ and policy parameter θ_0.
repeat
  Generate x_1^s
  for $i = 1$ to N **do**
    for $k = 1$ to T **do**
      Generate $l_k^{s,i}$ by sampling from $\pi_{\theta_s}\left(\cdot \mid x_k^{s,i}\right)$
      $x_{k+1}^{s,i} \leftarrow G\left(x_k, F_{l_k^{s,i}}\left(x_k^{s,i}\right), w_k^{s,i}\right)$
    end for
  end for
  Compute $\eta_{\theta_s}(\cdot,\cdot)$ by solving (9) using stochastic gradient descent method and sample trajectories $\left\{l_{1:T}^{s,1},\ldots,l_{1:T}^{s,N}\right\}$ and $\left\{x_{1:T+1}^{s,1},\ldots,x_{1:T+1}^{s,N}\right\}$
  $\Phi_{\theta_s}\left(x_{1:T+1}\right) \leftarrow \sum_k \eta_{\theta_s}\left(x_{k+1}, x_k\right)$
  Compute θ_{s+1} using (7)
  $s \leftarrow s+1$
until θ_s converges

Remark 4: In the case of nonlinear systems, one can show that $\mathsf{Pr}\left(L_k = j|\, x_{1:k+1}\right)$ is nonzero if there exists at least one w_k such that $(i) x_{k+1} = G\left(x_k, F_j(x_k), w_k\right)$, and (ii) the probability density function of W_k is nonzero everywhere.

V. Numerical Evaluations

In this section, we numerically study the performance of the proposed gain selection problem for different linear and nonlinear systems and evaluate the security-performance trade-off for these systems. In our simulations, for each system, the gain selection policy at each time step randomly selects a gain from a feedback gain collection $\{K_1,\ldots,K_M\}$ that includes the optimal feedback gain, *i.e.,* the feedback gain that minimizes the control performance in (2) when λ is zero.

TABLE I
The Parameters of the Linear Systems

Linear System 1	Linear System 2
$A = \begin{bmatrix} 1.98 & -0.98 \\ 1 & 0 \end{bmatrix}$	$A = \begin{bmatrix} 0.18 & -0.18 \\ -0.6 & 0 \end{bmatrix}$
$B = \begin{bmatrix} 0.0625 \\ 0 \end{bmatrix}$	$B = \begin{bmatrix} 1.56 \\ 0 \end{bmatrix}$
$K_1 = [3.27, -2.9]$	$K_1 = [0.0911, -0.5888]$
$K_2 = [2.95, -2.9]$	$K_2 = [0.0911, -0.8888]$
$K_3 = [3.27, -3.2]$	$K_3 = [0.3911, -0.0888]$

As the benchmark, we consider an open-loop gain selection policy that, at each time step, randomly selects a feedback gain from $\{K_1,\ldots,K_M\}$ irrespective of the system's state. More specifically, the open-loop gain selection policy selects the optimal gain in $\{K_1,\ldots,K_M\}$ with probability p and selects each remaining feedback gain with probability $\frac{1-p}{M-1}$. Thus, the open-loop gain selection policy is independent of the system's state. We refer to this policy as open-loop since it does not depend on the parameters of the closed-loop control system. We vary the parameter p to study the security-performance trade-off for each system under the open-loop gain selection policy.

To capture the inference ability of the adversary, we assume that the adversary uses a maximum likelihood estimator to infer the index of the feedback gain used by the system. We use the probability that the adversary correctly estimates the gain index as a measure of the adversary's inference ability. Note that the probability of correct inference effectively captures the security level of the system, *i.e.,* when the probability of correct inference is close to one, the system is less secure as the adversary can reliably infer the control gain. However, the system becomes more secure as the probability of correct inference is close to $1/M$ as the adversary has the maximum ambiguity about the selected feedback gain.

In what follows, we select the parameters of the linear and non-linear system such that the resulting systems are stabilizable. Moreover, the candidate gains set $\{K_1,\ldots,K_M\}$ for each system, includes the optimal feedback gain minimizing the quadratic control cost as well as slightly perturbed versions of the optimal feedback gain.

A. Linear Systems

In this section, we study the performance of the proposed gain selection policy for two different linear systems. The parameters A and B, as well as the collection of feedback gains for each system is shown in Table I. In what follows, we refer to these systems as Linear System 1 and Linear System 2. For each system, K_2 is the optimal gain and K_1, K_3 are perturbed versions of K_2. In our simulations, for each linear system, the process noise W_k is modeled as a sequence of independent and identically random vectors with zero mean and covariance matrix $\begin{bmatrix} 0.05 & 0 \\ 0 & 0.01 \end{bmatrix}$. We set $R = 1$ and $Q = 0.04 I_2$ where I_2 is a 2-by-2 identity matrix.

Fig. 3(a) and Fig. 3(b) show the control cost and the probability of correct inference by the adversary for Linear System 1 and Linear System 2, respectively, as a function

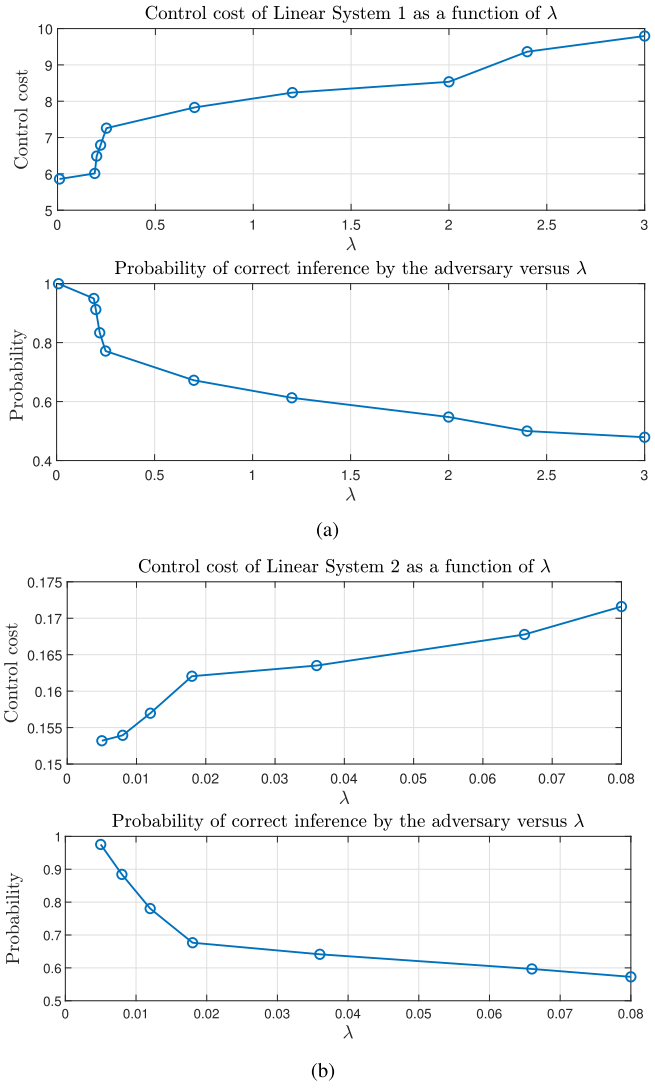


Fig. 3. The control cost and probability of correct inference by the adversary as a function of λ for Linear System 1 (*a*) and Linear System 2 (*b*).

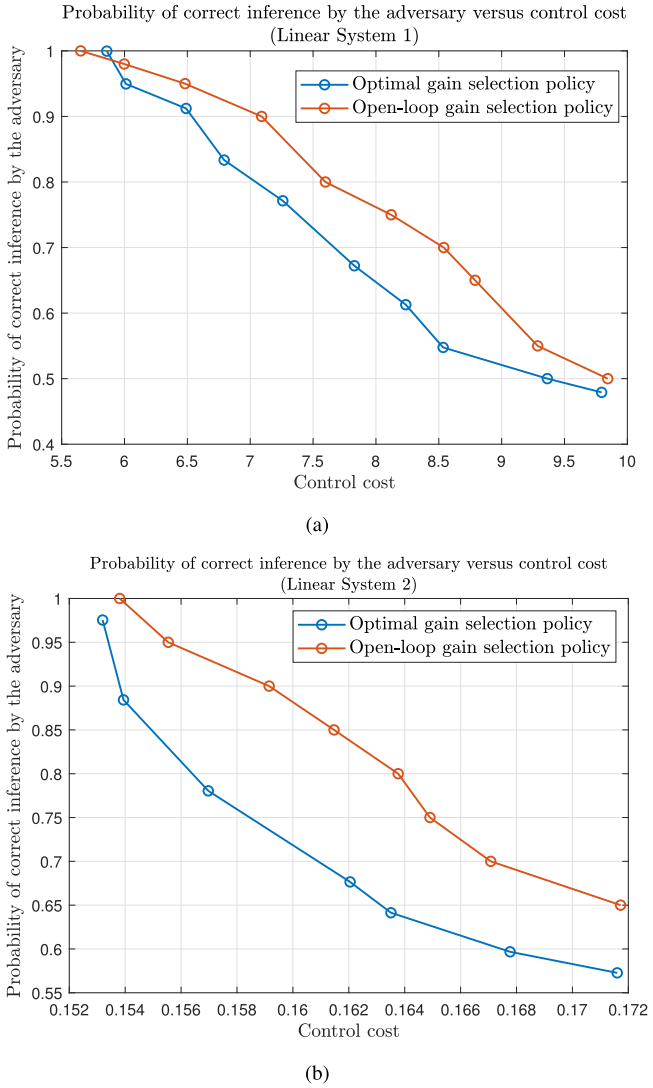


Fig. 4. The trade-off between the control cost and probability of correct inference by the adversary under the proposed gain selection policy and the open-loop gain selection policy for Linear System 1 (*a*) and Linear System 2 (*b*).

ofλ . When λ is equal to zero, the gain selection policy always selects the optimal feedback gain, *i.e.,* K_2. As a result, the adversary can accurately infer the feedback gain used by the system. Moreover, the control cost takes its smallest value when λ is equal to zero, as the optimal feedback gain is always selected by the gain selection policy. Note that K_2 minimizes the control cost of the system. As λ becomes large, the gain selection policy changes the feedback gain more often, and the feedback gains K_1 and K_3 are selected more frequently, which results in a higher level of ambiguity for the adversary. Consequently, the probability of correct inference by the adversary decreases as λ becomes large. According to Fig. 4(a) and Fig. 4(b), the control cost of the system increases as λ becomes large since the optimal feedback gain is selected less often, which results in a higher control cost.

Fig. 3(a) and Fig. 3(b) show the trade-off between the probability of correct inference by the adversary and the control cost for Linear System 1 and Linear System 2, respectively, under the optimal gain selection policy and the open-loop gain selection policy. To plot these figures, we varied λ and calculated the control cost and probability of correct inference under the optimal gain selection policy for each value of λ. We also varied the parameter p in the open-loop gain selection policy and calculated the corresponding control cost and probability of correct inference for each value of p. According to this figure, both the optimal and the open-loop gain selection policies ensure security at the expense of increasing the control cost. This is because, to decrease the probability of correct inference by the adversary, the non-optimal feedback gains K_1, K_3 must be selected more often that results in a higher control cost. However, under the optimal gain selection policy, a certain level of security, *i.e.,* the probability of correct inference, can be achieved at a lower control cost compared with the open-loop gain selection policy. This is because, under the open-loop policy, the feedback gains are randomly selected irrespective of the value of the system state. However, the optimal gain selection policy takes into account the value of the system states and the impact of the selected gain on the control cost.

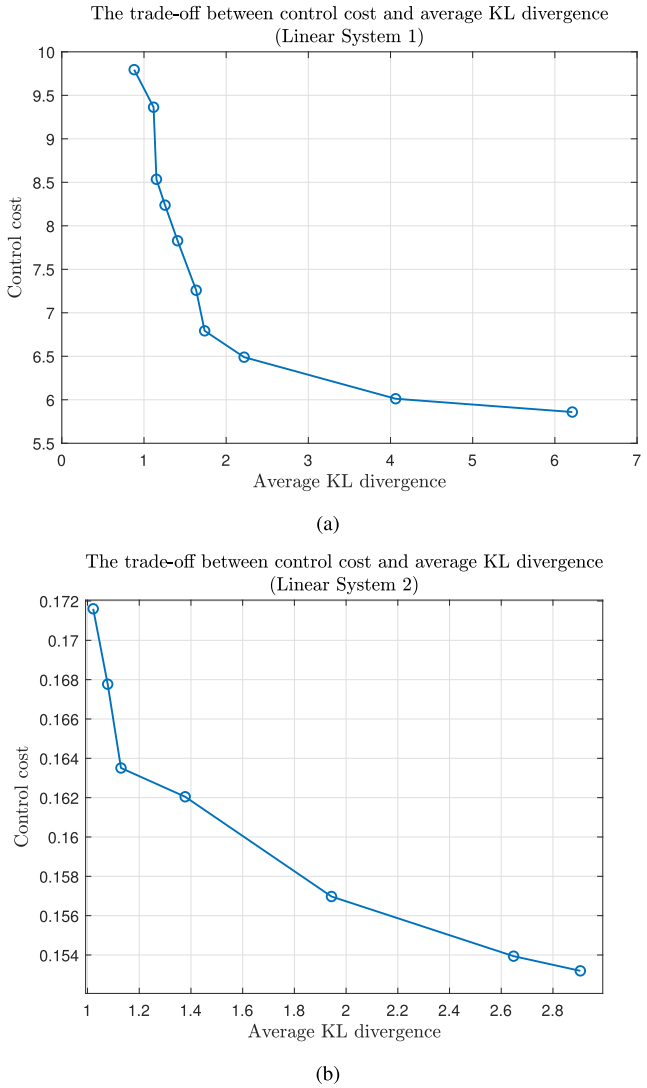


Fig. 5. The trade-off between the control cost and average KL divergence under the proposed gain selection policy for Linear System 1 (*a*) and Linear System 2 (*b*).

Fig. 5(a) and Fig. 5(b) show the trade-off between the control cost and the average KL divergence ($\mathsf{E}\left[\mathsf{D}\left[\mathsf{Uni}_M(j)\,\|\,\mathsf{Pr}\left(L_k = j|\,x_{1:k+1}\right)\right]\right]$) for Linear System 1 and Linear System 2, respectively. Note that a relatively large value of the average KL divergence indicates the feedback gains are easily distinguishable based on the system states. Thus, a large average KL divergence results in a low security level. However, a small average KL divergence indicates that the feedback gains are not easily distinguishable from the states. Thus, a small average KL divergence is desirable from the security perspective. Based on Fig. 5(a) and Fig. 5(b), the control cost increases as the KL divergence becomes small. This is because, at low values of the KL divergence, the optimal gain selection policy selects the feedback gains K_1 and K_3 more often to increase the ambiguous level of the adversary. However, the control cost increases at low values of KL divergence as the optimal gain K_2 is selected less often.

Finally, we study the temporal behavior of the optimal gain selection policy. Fig. 6 shows the sample trajectories of the

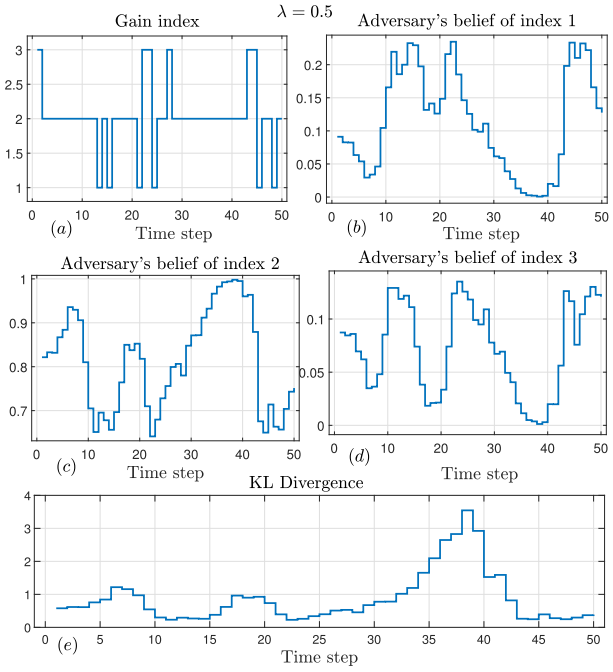


Fig. 6. The index of selected gain, the belief of adversary regarding each feedback gain, and the instantaneous value of the KL divergence for the Linear System 1 as a function time when λ is 0.5.

selected gain index, the belief of the adversary about each control gain, and the instantaneous value of the KL divergence ($\mathsf{D}\left[\mathsf{Uni}_M(j)\,\|\,\mathsf{Pr}\left(L_k = j|\,x_{1:k+1}\right)\right]$) for Linear System 1 when λ is equal to 0.5. Note that when λ is small, the optimization problem (2) assigns more weight to the control performance rather than the ambiguity of the adversary about the selected control gain. As a result, the gain selection policy selects K_2 with a high probability, and as a result, the index of selected gain is often equal to 2, as shown in Fig. 6-(a). When the gain selection policy chooses K_2 in a relatively long period, the belief of the adversary about K_2 becomes close to 1, as shown in Fig. 6-(b). As a result, the gain selection policy randomly selects a different gain in order to increase the uncertainty of the adversary about the system model when K_2 has been selected in a relatively long period.

Fig. 7 shows the sample trajectories of the gain index, the belief of the adversary, and the instantaneous value of the KL divergence for the Linear System 1 when λ is equal to 18.5. According to this figure, the gain selection policy is more random compared with $\lambda = 0.5$. Thus, the feedback is changed more often. Moreover, the adversary's belief about the system model is very low in this case due to the higher weight on the adversary's uncertainty level. Note that, under the optimal gain selection policy, the feedback gain is selected based on the value of the state, the instantaneous value of KL divergence, and the value of the optimal value (cost-to-go) function, *e.g.*, see equation (5). Thus, a policy that changes the gain only based on the instantaneous value of the KL divergence will not be optimal. The temporal behavior of the optimal gain selection policy for Linear System 2 is similar to that of Linear System 1 and is skipped to avoid repetition.

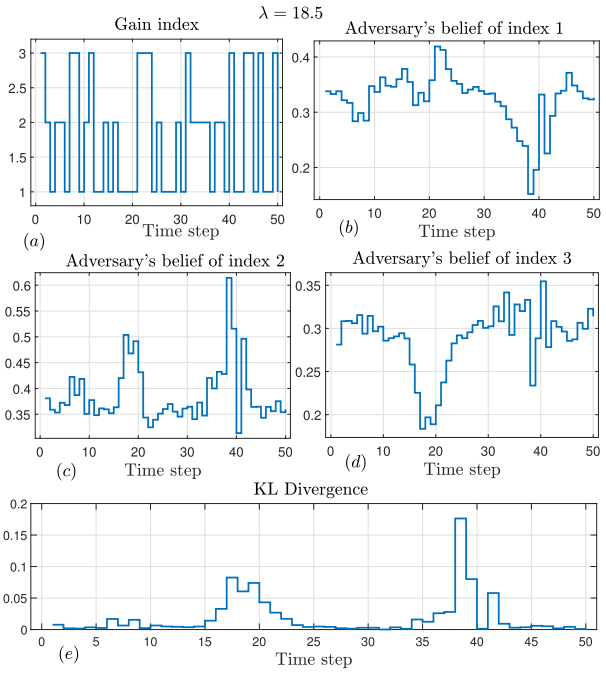


Fig. 7. The index of selected gain, the belief of adversary regarding each feedback gain, and the instantaneous value of the KL divergence for the Linear System 1 as a function time when λ is 18.5.

B. Nonlinear System

In the nonlinear case, we consider the following dynamical model

$$x_{k+1}^1 = x_k^1 + 0.01u_k^1 - 0.029\sin\left(x_k^1 - x_k^2\right) + 0.01w_k$$
$$x_{k+1}^2 = x_k^2 + 0.01u_k^2 - 0.029\sin\left(x_k^2 - x_k^1\right),$$

where x_k^1 and x_k^2 are the system states, u_k^1 and u_k^2 are the control inputs and $\{w_k\}_k$ is a sequence of iid Gaussian random variables with zero mean and unit variance. This model captures the frequency synchronization between two coupled oscillators, where x_k^1 and x_k^2 are the phases of two oscillators. Similar models have been used to study the frequency synchronization phenomenon in power systems [32]. In our simulations, we assume that the angular velocity of the second oscillator is 0.5, *i.e.,* $u_k^2 = 0.5$ and the first oscillator uses feedback to synchronize its frequency with the second oscillator. Also, the gain selection policy at each time step randomly selects one of the following feedback laws

$$F_1(x_k) = -20\left(x_k^1 - x_k^2\right) + 0.25,$$
$$F_2(x_k) = 5\left(x_k^1 - x_k^2\right) - 2.5,$$
$$F_3(x_k) = -5\left(x_k^1 - x_k^2\right) + 0.5,$$
$$F_4(x_k) = -6\left(x_k^1 - x_k^2\right) + 3,$$

where $F_2(\cdot)$ is the best control law for this system.

Fig. 8 shows the behavior of the control cost of the nonlinear system and the probability of correct inference by the adversary for this system versus λ. According to this figure, the control cost of the nonlinear system increases as λ becomes large since the optimal control law is selected less often when λ is large. Moreover, the probability of correct inference by the adversary decreases with λ as the gain selection policy changes the utilized control law more often.

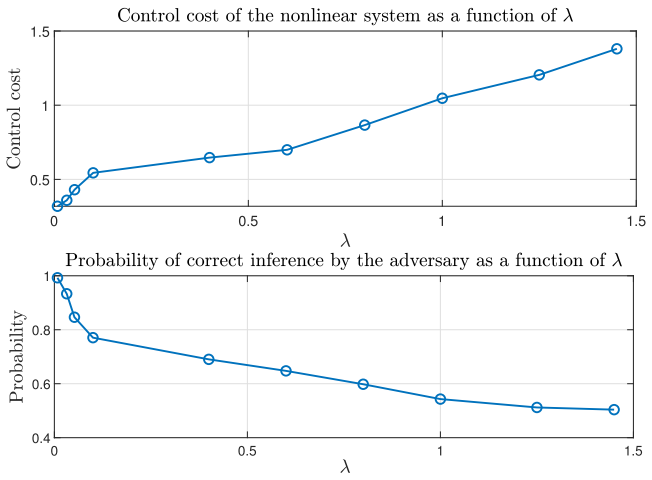


Fig. 8. The control cost and probability of correct inference by the adversary for the nonlinear system as a function of λ.

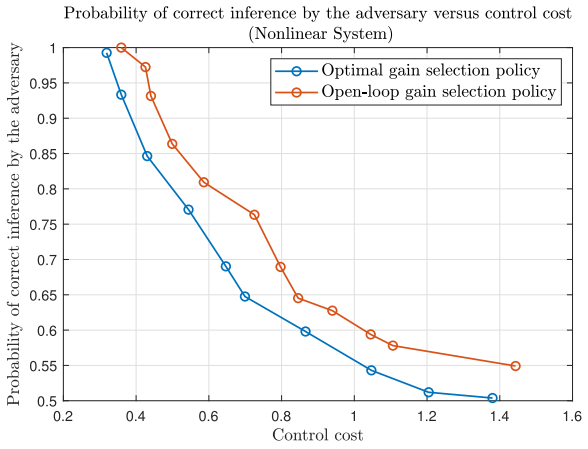


Fig. 9. The probability of correct inference by the adversary as a function of the control cost for the nonlinear system.

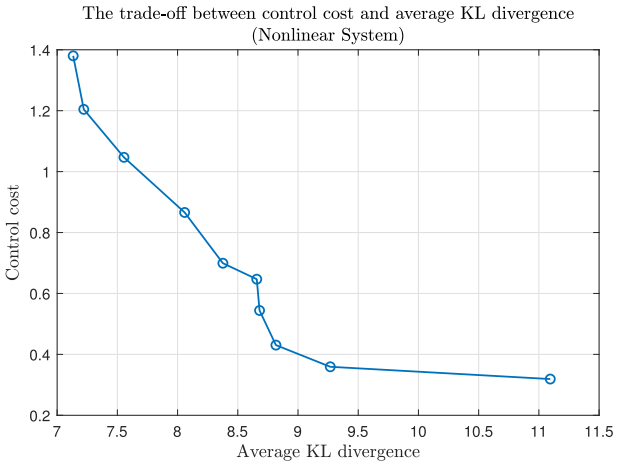
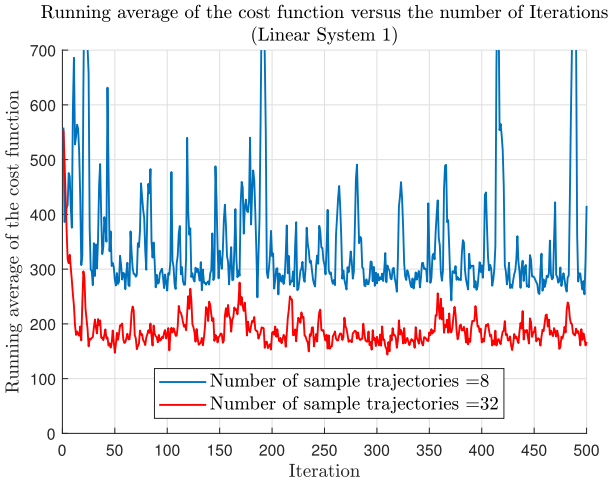
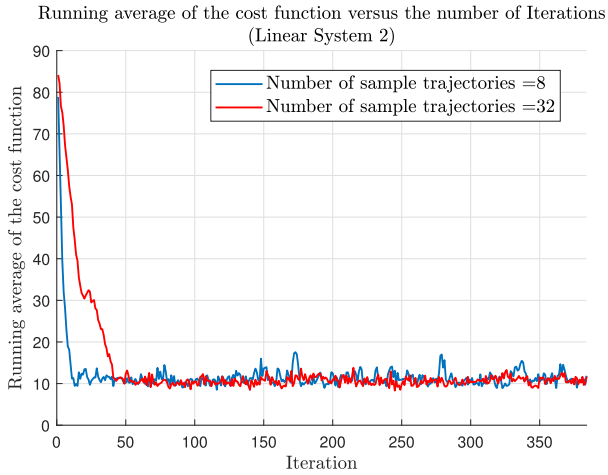


Fig. 10. The trade-off between the control cost and the average KL divergence under the proposed gain selection policy for the nonlinear system.

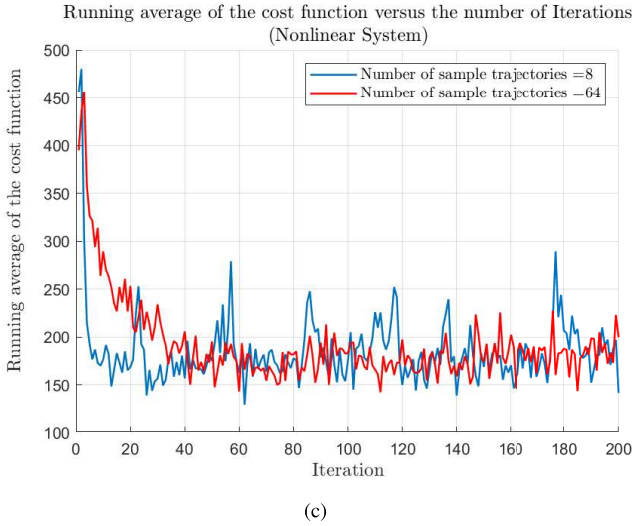
Fig. 9 shows the trade-off between the control cost and probability of the correct inference for the nonlinear system under the optimal gain selection policy and open-loop policy.



(a)



(b)



(c)

Fig. 11. The running average of the cost function of the optimization problem 2 under the policy gradient algorithm versus the number of iterations for the Linear System 1 (*a*), Linear System 2 (*b*), and the nonlinear system (*c*).

According to this figure, the optimal policy outperforms the open-loop policy. Moreover, the control cost increases as the probability of correct inference decreases. This is because the optimal control law is selected less often when the inference probability is low, which results in a higher control cost. Finally, Fig. 10 shows the trade-off between the control cost and the average KL divergence for the nonlinear system. Note that a large value of the average KL divergence indicates that the control laws are easily distinguishable by the adversary using the system state. To reduce the value of the KL divergence, the optimal gain selection policy changes the utilized control law more often, resulting in higher control costs as the optimal control law is selected less often.

C. Convergence of the Policy Gradient Algorithm

In this subsection, we numerically study the convergence behavior of the proposed policy gradient algorithm for Linear System 1, Linear System 2, and the nonlinear system. To this end, we computed the running average of the cost function of the optimization problem (2) at different iterations of the algorithm for each system with different numbers of sample trajectories N. The running average of the cost function at iteration s is defined as $\frac{1}{s}\sum_{i=1}^{s} f(\theta_s)$ where $f(\theta_s)$ is the value of the cost function in (2) when the policy parameter is equal to θ_s. Note that the asymptotic convergence of the running average indicates the convergence of the value of the cost function.

Fig. 11(a) shows the running average of the cost function under the policy gradient algorithm for Linear System 1 and different number of sample trajectories, *i.e., N*. According to Fig. 11(a), the running average of the cost function is highly random when N is equal to 8. Note that in the stochastic gradient algorithm, the true gradient is estimated using the sample trajectories. Thus, the estimate of the gradient is random. Therefore, the policy parameter and the value of the cost function may fluctuate in a wide range due to the randomness of the gradient estimate. However, as N becomes large, the estimate of the gradient is more accurate, resulting in small fluctuations in the running average of the cost function.

Fig. 11(b) and Fig. 11(c) show the behavior of the running average of the cost function with the number of iterations for Linear System 2 and the nonlinear system, respectively. Based on these figures, the fluctuations in the running average of the cost function decrease as the number of sample trajectories becomes large. Comparing Fig. 11(a) and Fig. 11(b), we observe that the behavior of the policy gradient algorithm under Linear System 1 is more random than that under Linear System 2. This is because the parameters of Linear System 1 are larger than those of Linear System 2. As a result, the impact of the process noise is amplified under Linear System 1.

VI. Conclusion and Future Work

In this paper, we developed a gain randomization framework against inference attacks on networked systems. In our setup, an adversary with access to the states of a system attempts to infer the system model. To prevent the inference attack, we developed a gain randomization framework in which the feedback control gain of the system is randomly changed at regular intervals. We formulated the gain selection problem as an optimal control problem with two objectives namely (*i*) reducing the control cost of the system while (*ii*) asymptotically maximizing the uncertainty level of the adversary about

the system model. We derived the Bellman optimality principle for the grain selection problem and studied the structural properties of the optimal gain selection policy. We also derived a policy gradient theorem, which allowed us to develop a stochastic gradient descent algorithm for computing an optimal policy. Our future work includes the optimal design of the gain selection policy in the presence of an active adversary that, in addition to overhearing the communication, can also influence the closed-loop behavior of the system, e.g., a false data injection attack scenario. In this case, the interaction between the security mechanism and the adversary can be modeled as a non-cooperative game where the adversary designs the additive signal to improve the detectability of the control gains, whereas the security mechanism designs the gain randomization policy to maximize the ambiguity level of the adversary. Another interesting direction for our future work includes the optimal design of the gain selection policy for continuous-time systems, with partial access to the state measurements and colored measurement and process noise.

APPENDIX A
PROOF OF THEOREM 1

The value function at time $T+1$ is given by

$$V^{\star}_{T+1}\left(x_{T+1}\right) = x^{\top}_{T+1} Q x_{T+1}.$$

Thus, the optimal value function at time $T+1$ only depends on x_{T+1}. Next, we show by induction that if the optimal value function at time $k+1$ only depends on x_{k+1}, then the value optimal value function at time k only depends on x_k. Note that the optimal value function at time k can be written as in (12), shown at the bottom of the page. Since x_k is in I_k, we have

$$\mathsf{E}\left[X^{\top}_k Q X_k \middle| I_k\right] = x^{\top}_k Q x_k. \tag{18}$$

Moreover, $\mathsf{E}\left[U^{\top}_k R U_k \middle| I_k\right]$ can be written as

$$\begin{aligned}\mathsf{E}\left[U^{\top}_k R U_k \middle| I_k\right] &\overset{(a)}{=} \mathsf{E}\left[X^{\top}_k L^{\top}_k R L_k X_k \middle| I_k\right] \\ &\overset{(b)}{=} x^{\top}_k \mathsf{E}\left[L^{\top}_k R L_k \middle| I_k\right] x_k\end{aligned}$$

$$\begin{aligned}V^{\star}_k(I_k) = \min_{\pi_k(\cdot|I_k)} \mathsf{E}\Big[& U^{\top}_k R U_k + X^{\top}_k Q X_k + \lambda \mathsf{D}\left[\mathsf{Unif}(M) \,\middle\|\, \mathsf{Pr}\left(L_k \middle| X_{1:k+1}\right)\right] + V^{\star}_{k+1}\left(X_{k+1}\right) \Big| I_k\Big] \\ & \times \mathsf{D}\left[\mathsf{Uni}_M(j) \,\middle\|\, \mathsf{Pr}\left(L_k = j \middle| x_{1:k+1}\right)\right]\end{aligned} \tag{12}$$

$$\begin{aligned}&= \frac{1}{M}\sum_j \log\left(\frac{1/M}{\mathsf{Pr}\left(L_k = j \middle| x_{1:k+1}\right)}\right) \\ &= -\log(M) - \frac{1}{M}\sum_j \log\left(\mathsf{Pr}\left(L_k = j \middle| x_{1:k+1}\right)\right)\end{aligned} \tag{13}$$

$$\begin{aligned}\mathsf{Pr}\left(L_k = j \middle| x_{1:k+1}\right) &\overset{(a)}{=} \frac{p\left(x_{k+1}|x_k, L_k = j\right)\mathsf{Pr}\left(L_k = j \middle| x_{1:k}\right)p\left(x_{1:k}\right)}{p\left(x_{1:k+1}\right)} \\ &\overset{(b)}{=} \frac{p\left(x_{k+1}|x_k, L_k = j\right)\mathsf{Pr}\left(L_k = j \middle| x_{1:k}\right)p\left(x_{1:k}\right)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\mathsf{Pr}\left(L_k = i \middle| x_{1:k}\right)p\left(x_{1:k}\right)} \\ &\overset{(c)}{=} \frac{p\left(x_{k+1}|x_k, L_k = j\right)\pi_k\left(j|I_k\right)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_k\left(i|I_k\right)}\end{aligned} \tag{14}$$

$$\begin{aligned}&\mathsf{E}\left[\mathsf{D}\left[\mathsf{Uni}_M(j) \,\middle\|\, \mathsf{Pr}\left(L_k = j \middle| X_{1:k+1}\right)\right] \middle| I_k\right] \\ &= -\log(M) - \frac{1}{M}\sum_j \mathsf{E}\left[\log\frac{p\left(X_{k+1}|x_k, L_k = j\right)\pi_k\left(j|I_k\right)}{\sum_i p\left(X_{k+1}|x_k, L_k = i\right)\pi_k\left(i|I_k\right)}\middle| I_k\right]. \\ &= -\log(M) - \frac{1}{M}\sum_j \int \log\frac{p\left(x_{k+1}|x_k, L_k = j\right)\pi_k\left(j|I_k\right)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_k\left(i|I_k\right)} p\left(x_{k+1}|I_k\right)dx_{k+1} \\ &\overset{(a)}{=} -\log(M) - \frac{1}{M}\sum_{j,l} \int \log\frac{p\left(x_{k+1}|x_k, L_k = j\right)\pi_k\left(j|I_k\right)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_k\left(i|I_k\right)} p\left(x_{k+1}|x_k, L_k = l\right)\pi_k\left(l|I_k\right)dx_{k+1}\end{aligned} \tag{15}$$

$$\begin{aligned}\mathsf{E}\left[V^{\star}_{k+1}\left(X_{k+1}\right)\middle| I_k\right] &= \int V^{\star}_{k+1}\left(x_{k+1}\right)p\left(x_{k+1}|I_k\right)d_{x_{k+1}} \\ &= \sum_l \int V^{\star}_{k+1}\left(x_{k+1}\right)p\left(x_{k+1}|x_k, L_k = l\right)\pi_k\left(l|I_k\right)dx_{k+1}\end{aligned} \tag{16}$$

$$\begin{aligned}V^{\star}_k(I_k) = \min_{\pi_k(\cdot|I_k)} & x^{\top}_k Q x_k + \sum_l x^{\top}_k K^{\top}_l R K_l x_k \pi_k\left(l|I_k\right) + \log\frac{1}{M} \\ & + \frac{1}{M}\sum_{j,l}\int V^{\star}_{k+1}\left(x_{k+1}\right) - \lambda\log\frac{p\left(x_{k+1}|x_k, L_k = j\right)\pi_k\left(j|I_k\right)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_k\left(i|I_k\right)} p\left(x_{k+1}|x_k, L_k = l\right)\pi_k\left(l|I_k\right)dx_{k+1}\end{aligned} \tag{17}$$

$$\overset{(c)}{=} x_k^\top \sum_l K_l^\top R K_l \mathsf{Pr}\left(L_k = l | I_k\right) x_k$$
$$\overset{(d)}{=} \sum_l x_k^\top K_l^\top R K_l x_k \pi_k\left(l | I_k\right), \tag{19}$$

where (a) follows from the definition of U_k, (b) follows from the fact that x_k is in I_k, (c) and (d) follow from the definition of the conditional expectation and the definition of the policy, respectively.

Using the definition of the KL divergence, we have as in (13), shown at the bottom of the previous page. Using the Bayes' rule, $\mathsf{Pr}\left(L_k = j | x_{1:k+1}\right)$ can be expanded as in (14), shown at the bottom of the previous page, where (a) and (b) follow from the following Markov chain $X_{1:k-1} \longrightarrow (X_k, L_k) \longrightarrow X_{k+1}$, and (c) follows from the definition of the policy. Combining (14) and (13), we have as in (15), shown at the bottom of the previous page, where (a) follows from the Bayes' rule and the Markov chain $X_{1:k-1} \longrightarrow (X_k, L_k) \longrightarrow X_{k+1}$. Finally, $\mathsf{E}\left[V_{k+1}^\star\left(X_{k+1}\right) \middle| I_k\right]$ can be written as in (16), shown at the bottom of the previous page. Combining (12)-(16), we have as in (17), shown at the bottom of the previous page.

Note that the objective function of the optimization above only depends on x_k, the gain selection policy and system parameters. Then, the optimal gain selection policy $\pi_k^\star(\cdot | I_k)$ is only a function of x_k, and it does not depend on the entire history of the state measurements. Moreover, the optimal value function at time k only depends on x_k.

To show that (2) is a nonlinear MDP, note that the dynamic of the system in (1) is Markovian. Using (13)-(14), the per stage cost of the (2) at time $k < T$ can be written as

$$x_k^\top Q x_k + x_k^\top K_l^\top R K_l x_k + \log \frac{1}{M} - \lambda \log \frac{p\left(x_{k+1} | x_k, L_k = j\right)\pi_k\left(j | I_k\right)}{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\pi_k\left(i | I_k\right)},$$

which only depends on X_k, X_{k+1}, L_k. Also, the per stage cost is nonlinear in the gain selection policy. Thus, the optimization problem (2) is a nonlinear MDP.

$$\begin{aligned}
\nabla_\theta O_1^\theta &= \nabla_\theta \sum_{l_{1:T}} \int r\left(x_{1:T+1}, l_{1:T}\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&= \sum_{l_{1:T}} \int r\left(x_{1:T+1}, l_{1:T}\right) \left(\nabla_\theta \log \prod_{k=1}^{T} \pi_\theta\left(l_k | x_k\right)\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&= \sum_{l_{1:T}} \int r\left(x_{1:T+1}, l_{1:T}\right) \left(\sum_{k=1}^{T} \nabla_\theta \log \pi_\theta\left(l_k | x_k\right)\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&= \mathsf{E}\left[\left(X_{T+1}^\top Q X_{T+1} + \sum_{k=1}^{T} X_k^\top Q X_k + U_k^\top R U_k\right) \sum_{k=1}^{T} \nabla_\theta \log \pi_\theta\left(L_k | X_k\right)\right]
\end{aligned} \tag{20}$$

$$\begin{aligned}
\nabla_\theta O_2^\theta &= \nabla_\theta \sum_{l_{1:T}} \int \Phi_\theta\left(x_{1:T+1}\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&= \sum_{l_{1:T}} \int \nabla_\theta \Phi_\theta\left(x_{1:T+1}\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&\quad + \sum_{l_{1:T}} \int \Phi_\theta\left(x_{1:T+1}\right) \left(\nabla_\theta \log \prod_{k=1}^{T} \pi_\theta\left(l_k | x_k\right)\right) p_1\left(x_1\right) \prod_{k=1}^{T} p\left(x_{k+1} | x_k, L_k = l_k\right)\pi_\theta\left(l_k | x_k\right) dx_{1:T+1} \\
&= \mathsf{E}\left[\nabla_\theta \Phi_\theta\left(X_{1:T+1}\right) + \Phi_\theta\left(X_{1:T+1}\right) \left(\nabla_\theta \log \prod_{k=1}^{T} \pi_\theta\left(L_k | X_k\right)\right)\right]
\end{aligned} \tag{21}$$

$$\Phi_\theta\left(x_{1:T+1}\right) = \frac{1}{M} \sum_{k=1}^{T} \sum_{j=1}^{M} \log \frac{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\pi_\theta\left(i | x_k\right)}{p\left(x_{k+1} | x_k, L_k = j\right)\pi_\theta\left(j | x_k\right)} \tag{22}$$

$$\begin{aligned}
\nabla_\theta \Phi_\theta\left(x_{1:T+1}\right) &= \frac{1}{M} \sum_{k=1}^{T} \sum_{j=1}^{M} \frac{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\nabla_\theta \pi_\theta\left(i | x_k\right)}{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\pi_\theta\left(i | x_k\right)} - \frac{\nabla_\theta \pi_\theta\left(j | x_k\right)}{\pi_\theta\left(j | x_k\right)}. \\
&= \sum_{k=1}^{T} \frac{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\nabla_\theta \pi_\theta\left(i | x_k\right)}{\sum_i p\left(x_{k+1} | x_k, L_k = i\right)\pi_\theta\left(i | x_k\right)} - \frac{1}{M} \sum_{j=1}^{M} \frac{\nabla_\theta \pi_\theta\left(j | x_k\right)}{\pi_\theta\left(j | x_k\right)}
\end{aligned} \tag{23}$$

To show convexity, note that $\mathsf{E}\left[U_k^\top R U_k \middle| I_k\right]$ and $\mathsf{E}\left[V_{k+1}^\star\left(X_{k+1}\right)\middle| I_k\right]$ are linear in the policy (see (19) and (16)). Let $z = p\left(x_{k+1}|x_k, L_k = j\right)\pi_k(j|I_k)$ and $t = \sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_k(i|I_k)$. Then, $t\log\frac{z}{t}$ is concave since $\log(z)$ is concave and $t\log\frac{z}{t}$ is the perspective of $\log(z)$. Thus, the integrand in (15) is concave in the policy, which implies that the integral in (15) is a concave function of the policy since the sum of concave functions is concave. Finally, $\mathsf{E}\left[\mathsf{D}\left[\mathsf{Uni}_M(j)\middle\|\mathsf{Pr}\left(L_k = j|X_{1:k+1}\right)\right]\middle| I_k\right]$ is convex in the policy due to the negative sign outside the integral in (15).

Appendix B
Proof of Theorem 2

Note that the objective function of the optimization problem (2) can be written as $O_1^\theta + \lambda O_2^\theta$ where

$$O_1^\theta = \mathsf{E}\left[X_{T+1}^\top Q X_{T+1} + \sum_{k=1}^{T} X_k^\top Q X_k + K_{L_k}^\top X_k^\top R K_{L_k} X_k\right],$$

$$O_2^\theta = \mathsf{E}\left[\sum_{k=1}^{T}\mathsf{D}\left[\mathsf{Unif}(M)\middle\|\mathsf{Pr}\left(L_k|X_{1:k+1}\right)\right]\right].$$

We first derive an expression for $\nabla_\theta O_1^\theta$ as in (21), shown at the bottom of the previous page, where $r\left(x_{1:T+1}, l_{1:T}\right) = x_{T+1}^\top Q x_{T+1} + \sum_{k=1}^{T} x_k^\top Q x_k + K_{l_k}^\top x_k^\top R K_{l_k} x_k$, and $l_{1:T}$ is a realization of $L_{1:T} = \{L_1, \ldots, L_T\}$.

Moreover, $\nabla_\theta O_2^\theta$ can be written as (21) where $\Phi_\theta\left(x_{1:T+1}\right)$ is given by as in (22), shown at the bottom of the previous page.

Note that $\nabla_\theta\Phi_\theta\left(x_{1:T+1}\right)$ can be expanded as in (23), shown at the bottom of the previous page.

We also have as in (24), shown at the bottom of the page.

Thus, $\nabla_\theta O_2^\theta$ can be written as in (25), shown at the bottom of the page.

Appendix C
Proof of Theorem 3

To prove this result, we first derive an expression for the function $\eta_\theta\left(x_{k+1}, x_k\right)$ in terms of the KL divergence. Note that using (13) and (14), we have

$$\eta_\theta\left(x_{k+1}, x_k\right) = \log(M) + \mathsf{D}\left[\mathsf{Uni}_M(j)\middle\|\mathsf{Pr}\left(L_k = j|x_{1:k+1}\right)\right]. \tag{28}$$

We next show that the KL divergence term in (28) can be expressed as the solution of an optimization problem. To this end, let $\tilde{L}$ to denote a uniformly distributed random variable over $\{1, \ldots, M\}$ which is independent of $\left\{X_{1:T+1}, L_{1:T}\right\}$. Then, we have as in (26), shown at the bottom of the page, where

$$\begin{aligned}
\mathsf{E}\left[\nabla_\theta\Phi_\theta\left(X_{1:T+1}\right)\right] &= \mathsf{E}\left[\frac{\sum_i p\left(X_{k+1}|X_k, L_k = i\right)\nabla_\theta\pi_\theta(i|X_k)}{\sum_i p\left(X_{k+1}|X_k, L_k = i\right)\pi_\theta(i|X_k)}\right] \\
&= \int\frac{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\nabla_\theta\pi_\theta(i|x_k)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_\theta(i|x_k)} p_\theta\left(x_{k+1}, x_k\right) dx_{k+1}dx_k \\
&= \int\frac{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\nabla_\theta\pi_\theta(i|x_k)}{\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_\theta(i|x_k)}\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_\theta(i|x_k)p(x_k)\,dx_{k+1}dx_k \\
&= \nabla_\theta\int\sum_i p\left(x_{k+1}|x_k, L_k = i\right)\pi_\theta(i|x_k)p(x_k)\,dx_{k+1}dx_k \\
&= \nabla_\theta\int p_\theta\left(x_{k+1}, x_k\right)dx_{k+1}dx_k = 0
\end{aligned} \tag{24}$$

$$\nabla_\theta O_2^\theta = \mathsf{E}\left[\Phi_\theta\left(X_{1:T+1}\right)\sum_{k=1}^{T}\nabla_\theta\log\pi_\theta(L_k|X_k) - \frac{1}{M}\sum_{k=1}^{T}\sum_{j=1}^{M}\nabla_\theta\log\pi_\theta(j|X_k)\right] \tag{25}$$

$$\begin{aligned}
\mathsf{E}\left[\mathsf{D}\left[\mathsf{Uni}_M(j)\middle\|\mathsf{Pr}\left(L_k = j|X_{1:k+1}\right)\right]\right] &\overset{(a)}{=} \mathsf{E}\left[\mathsf{D}\left[\mathsf{Uni}_M(j)\middle\|\mathsf{Pr}\left(L_k = j|X_k, X_{k+1}\right)\right]\right] \\
&= \int\sum_j\log\frac{\mathsf{Uni}_M(j)\,p\left(x_k, x_{k+1}\right)}{p\left(x_{k+1}|x_k, L_k = j\right)\pi_\theta(L_k = j|x_k)p(x_k)}\mathsf{Uni}_M(j)\,p\left(x_k, x_{k+1}\right)dx_kdx_{k+1} \\
&\overset{(b)}{=} \mathsf{D}\left[\mathsf{Uni}_M(j)\,p\left(x_k, x_{k+1}\right)\middle\|p\left(x_{k+1}|x_k, L_k = j\right)\pi_\theta(L_k = j|x_k)p(x_k)\right] \\
&\overset{(c)}{=} \sup_{f(\cdot,\cdot,\cdot)\in L_+^3}\mathsf{E}\left[\log f\left(\tilde{L}, X_k, X_{k+1}\right)\right] - \mathsf{E}\left[f\left(L_k, X_k, X_{k+1}\right)\right] + 1
\end{aligned} \tag{26}$$

$$\begin{aligned}
\mathsf{D}\left[\mathsf{Uni}_M(j)\middle\|\mathsf{Pr}\left(L_k = j|x_{1:k+1}\right)\right] &= \frac{1}{M}\sum_j\log\frac{\mathsf{Uni}_M(j)}{\mathsf{Pr}\left(L_k = j|x_k, x_{x+1}\right)}, \\
&\overset{(a)}{=} \frac{1}{M}\sum_j\log f^\star\left(j, x_k, x_{k+1}\right)
\end{aligned} \tag{27}$$

(a) follow from (14), (b) follows from the definition of the KL divergence, $p\left(x_k, x_{k+1}\right)$ is the joint density of X_k, X_{k+1}, $p\left(x_{k+1}\,|x_k, L_k = j\right)$ is the conditional density of X_{k+1} given $X_k = x_k, L_k = j$, and $p\left(x_k\right)$ is the density of X_k. In (26), (c) follows from a variational representation of the KL divergence which was developed in [33]. For the sake of completeness, we next state the aforementioned variational formulation. Let $\mathsf{P}\left(x\right)$ and $\mathsf{Q}\left(x\right)$ be two probability density functions. Then, the KL divergence between $\mathsf{P}\left(x\right)$ and $\mathsf{Q}\left(x\right)$ can be written as

$$\mathsf{D}\left[\mathsf{P}\left(x\right)\|\mathsf{Q}\left(x\right)\right] = \sup_{f(\cdot)>0} \mathsf{E}_{\mathsf{P}}\left[\log f(X)\right] - \mathsf{E}_{\mathsf{Q}}\left[f(X)\right] + 1.$$

Moreover, the supremum is attained at $f^{\star}\left(x\right) = \frac{\mathsf{P}(x)}{\mathsf{Q}(x)}$ [33]. Thus, the optimal solution of the optimization problem in (26), denoted as $f^{\star}\left(\cdot,\cdot,\cdot\right)$, is given by

$$\begin{aligned} &f^{\star}\left(j, x_k, x_{k+1}\right) \\ &\quad = \frac{\mathsf{Uni}_M\left(j\right) p\left(x_k, x_{k+1}\right)}{p\left(x_{k+1}\,|x_k, L_k = j\right)\pi_\theta\left(L_k = j|\, x_k\right) p\left(x_k\right)} \\ &\quad = \frac{\mathsf{Uni}_M\left(j\right)}{\mathsf{Pr}\left(L_k = j|\, x_k, x_{x+1}\right)}. \end{aligned} \tag{29}$$

Using the definition of KL divergence, we have as in (27), shown at the bottom of the previous page, where (a) follows from (29). Combining (28) and (27), we have

$$\eta_\theta\left(x_{k+1}, x_k\right) = \log\left(M\right) + \frac{1}{M}\sum_j \log f^{\star}\left(j, x_k, x_{k+1}\right),$$

which completes the proof. Note that, in the literature, the variational formulation is typically used as an efficient method of computing the KL divergence. However, in this theorem, we are interested in the optimal solution of the variational formulation, *i.e.,* $f^{\star}\left(j, x_k, x_{k+1}\right)$, as it provides an alternative way for computing $\eta_\theta\left(x_{k+1}, x_k\right)$. Recall that the direct computation of $\eta_\theta\left(x_{k+1}, x_k\right)$ is challenging in the nonlinear case.

REFERENCES

[1] S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of CPS security," *Annu. Rev. Control*, vol. 47, pp. 394–411, Jul. 2019.

[2] L. Bilge and T. Dumitraş, "Before we knew it: An empirical study of zero-day attacks in the real world," in *Proc. ACM Conf. Comput. Commun. Secur.*, Oct. 2012, pp. 833–844.

[3] M. B. Salem, S. Hershkop, and S. J. Stolfo, "A survey of insider attack detection research," in *Insider Attack and Cyber Security: Beyond the Hacker*. New York, NY, USA: Springer, 2008, pp. 69–90.

[4] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, Nov. 2013.

[5] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 3, pp. 636–653, May 2010.

[6] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[7] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," in *Proc. IEEE*, Sep. 2007, vol. 95, no. 1, pp. 138–162.

[8] C. N. Hadjicostis and R. Touri, "Feedback control utilizing packet dropping network links," in *Proc. 41st IEEE Conf. Decis. Control*, vol. 2, May 2002, pp. 1205–1210.

[9] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1454–1467, Jun. 2014.

[10] E. Chow and A. Willsky, "Analytical redundancy and the design of robust failure detection systems," *IEEE Trans. Autom. Control*, vol. AC-29, no. 7, pp. 603–614, Jul. 1984.

[11] J. R. Sklaroff, "Redundancy management technique for space shuttle computers," *IBM J. Res. Develop.*, vol. 20, no. 1, pp. 20–28, Jan. 1976.

[12] W. Abbas, A. Laszka, and X. Koutsoukos, "Improving network connectivity and robustness using trusted nodes with application to resilient consensus," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 4, pp. 2036–2048, Dec. 2018.

[13] M. Momani and S. Challa, "Survey of trust models in different network domains," 2010, *arXiv:1010.0168*.

[14] J. S. Baras and X. Liu, "Trust is the cure to distributed consensus with adversaries," in *Proc. 27th Medit. Conf. Control Autom. (MED)*, Jul. 2019, pp. 195–202.

[15] W. P. M. H. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *Proc. IEEE 51st IEEE Conf. Decis. Control (CDC)*, Dec. 2012, pp. 3270–3285.

[16] C. De Persis and P. Tesi, "On resilient control of nonlinear systems under denial-of-service," in *Proc. 53rd IEEE Conf. Decis. Control*, Dec. 2014, pp. 5254–5259.

[17] M. Pirani, A. Mitra, and S. Sundaram, "Graph-theoretic approaches for analyzing the resilience of distributed control systems: A tutorial and survey," *Automatica*, vol. 157, Nov. 2023, Art. no. 111264.

[18] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[19] F. Farokhi, I. Shames, and N. Batterham, "Secure and private control using semi-homomorphic encryption," *Control Eng. Pract.*, vol. 67, pp. 13–20, Oct. 2017.

[20] C. Dwork, "Differential privacy: A survey of results," in *Proc. Int. Conf. Theory Appl. Models Comput. (TAMC)*, 2008, pp. 1–19.

[21] J. Le Ny and G. J. Pappas, "Differentially private filtering," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 341–354, Feb. 2014.

[22] R. Motwani and P. Raghavan, "Randomized algorithms," *ACM Comput. Surveys*, vol. 28, no. 1, pp. 33–37, 1996.

[23] S. Weerakkody and B. Sinopoli, "Detecting integrity attacks on control systems using a moving target approach," in *Proc. 54th IEEE Conf. Decis. Control*, Dec. 2015, pp. 5820–5826.

[24] 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, Sep. 2019.

[25] E. T. M. Beltrán et al., "Mitigating communications threats in decentralized federated learning through moving target defense," *Wireless Netw.*, vol. 30, no. 9, pp. 7407–7421, Dec. 2024.

[26] E. Nowroozi, M. Mohammadi, P. Golmohammadi, Y. Mekdad, M. Conti, and S. Uluagac, "Resisting deep learning models against adversarial attack transferability via feature randomization," *IEEE Trans. Services Comput.*, vol. 17, no. 1, pp. 18–29, Jan. 2024.

[27] E. Nekouei, H. Sandberg, M. Skoglund, and K. H. Johansson, "A model randomization approach to statistical parameter privacy," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 839–850, Feb. 2023.

[28] E. Nekouei, M. Pirani, H. Sandberg, and K. H. Johansson, "A randomized filtering strategy against inference attacks on active steering control systems," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 16–27, 2022.

[29] F. Kato, Y. Cao, and M. Yoshikawa, "Preventing manipulation attack in local differential privacy using verifiable randomization mechanism," in *Proc. IFIP Annu. Conf. Data Appl. Secur. Privacy*. Cham, Switzerland: Springer, Sep. 2021, pp. 43–60, doi: 10.1007/978-3-030-81242-3_3.

[30] J. Zeng, J. Xu, X. Zheng, and X. Huang, "Certified robustness to text adversarial attacks by randomized," *Comput. Linguistics*, vol. 49, no. 2, pp. 395–427, Jun. 2023.

[31] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.

[32] F. Dörfler and F. Bullo, "Synchronization and transient stability in power networks and nonuniform Kuramoto oscillators," *SIAM J. Control Optim.*, vol. 50, no. 3, pp. 1616–1642, Jan. 2012.

[33] X. Nguyen, M. J. Wainwright, and M. I. Jordan, "Estimating divergence functionals and the likelihood ratio by convex risk minimization," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5847–5861, Nov. 2010.