This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3229774

Co-design for Resilience and Performance

Navid Hashemi, Justin Ruths

***Abstract*—We present two optimization approaches to minimize the impact of sensor falsification attacks in linear time invariant systems controlled by estimate-based feedback. We accomplish this by finding observer and controller gain matrices that minimize outer ellipsoidal bounds that contain the reachable set of attack-induced states. To avoid trivial solutions, we involve a covariance-based $\|H\|_2$ closed-loop performance constraint. This exposes a trade-off between system security and closed-loop performance and demonstrates that only small concessions in performance can lead to large gains in our reachability-based security metric. We provide both a nonlinear optimization based on geometric sums and a fully convexified approach formulated with linear matrix inequalities. We demonstrate the effectiveness of these tools on two numerical case studies.**

***Index Terms*—LMI, reachable set, security, robust control.**

I. Introduction

A growing awareness of security concerns in automated physical processes has increased interest in our ability to quantify the impact of would-be attackers. It is the networked nature and cyber portion of these modern control systems that provides improvements in efficiency, performance, and reliability as well as exposes new avenues for disruption through attack. It is also this large and distributed nature of cyber physical systems that requires a systematic approach for detecting anomalies rather than relying on operators to spot abnormal behavior. Work along these lines imposes a detector to raise alerts when sensor measurements do not fall in line with model-based predictions, thus constraining what an attacker can do while staying undetected [1]-[4]. The tuning of such detectors is a balancing act between increasing the sensitivity to attacks while reducing the number of false alarms (alerts raised during normal operation) [5]. The next clear step in this direction of research is to then minimize this impact through control system design. Initially work used the distance (norm) that an attacker could drive the state as a proxy for impact [5], [6], however, ultimately the reachable set provides important information about which components of a system are effected more than others and can inform whether the attacked state might reach dangerous regions of state space [7]. The quantification, or *analysis*, studies have often used ellipsoidal bounds on the actual attack-induced reachable set achieved either through iterative methods [1] or the satisfaction of linear matrix inequalities (LMIs) [3], [7]. The latter extends gracefully to optimization to address the *design* question, although doing so often requires re-linearizing the inequalities with respect to the new design variables.

In this paper, we leverage past work [3] on quantification of ellipsoidal bounds on the attack-induced reachable set to design the observer and controller gain matrices to minimize the ellipsoidal bound (and thus the reachable set) when a linear time invariant system is controlled with estimate-based feedback. As has been pointed out (see, e.g., [7]), it is important to pair a security minimization with a constraint on closed-loop performance, otherwise a trivial solution exists to disconnect the feedback loop and thus cut off the effect of the attack on the system state. Our design solutions then characterize a trade-off between resilience to attacks and closed-loop performance. Here we specify an input covariance-based (ICB) $\|H\|_2$ performance [8], which takes into account the covariance of the noise, unlike other distribution-agnostic approaches to robust control [9], [10].

The notion of a trade-off between security and robustness exists within the literature on Resilient Control [11], [12], which emphasizes that a system can be fragile and weak against cyber perturbations despite being robust against physical uncertainties [13], [14]. In this context, *robustness* refers to the performance of the system under physical uncertainties and disturbances and *security* refers to the safety of the system against malicious signals and uncertainties entering from the cyber layer. Recent work in this theme has primarily sought to identify criteria to avoid making a compromise between robustness and security, i.e., maintain uncorrupted state estimation or control over an attacked system [15]-[20]. The alternative is to quantify the system degradation possible due to worst-case attacks that are only constrained to avoid detection [3], [6], [21]. Our contribution here is to design the system to minimize this worst-case degradation, through optimization over bounds with some linearization assumptions.

Our work is distinguished from [7]by (a) using estimate feedback, which would create nonlinearities in the formulation in [7]; (b) approximating the residual covariance, which allows the optimization to maintain a fixed false alarm rate for the detector despite changes in the gain matrices; (c) using a covariance-informed $\|H\|_2$ performance metric as opposed to a distributionally insensitive version (we develop a novel convexification of the ICB $\|H\|_2$ design problem, by dividing the problem into a convex optimization and a generalized algebraic Riccati equation); (d) introducing a magnification factor to scale the shape matrices associated with the $\|H\|_2$ and reachable set decision variables; and (e) develop a nonlinear optimization based on geometric sums to more accurate probe the security-performance trade-off. Our tools enable us to demonstrate that small reductions in closed-loop performance can enable substantive improvements in our security metric when attackers seek to remain undetected.

$\mathbb{S}^n_{++}$ denotes the space of $n \times n$ positive definite matrices. We use $\overline{i,j} := i, i+1, \ldots, j-1, j$ to denote an interval of integers. J^{ij}_A is a matrix of all zeros with the same size as matrix A whose (i,j) entry is a one ($[J^{ij}_A]_{ij} = 1$).

The authors are with the Department of Mechanical Engineering at the University of Texas at Dallas. Correspondence to jruths@utdallas.edu. This material is based upon work supported by the National Science Foundation under Grant No. CMMI-2143485.

II. Background

We are motivated by modern networked control systems in which sensing, control, and actuation often occur in physically disparate locations and are coordinated through a wired and/or wireless communication network - such as in power distribution, where several hierarchical layers coordinate across large geographic regions; in process control, where plants are composed of many interacting but mostly siloed modules; in autonomous driving, where vehicles might combine local sensing and vehicle-to-vehicle communication in order to navigate. We model the physical dynamics with a discrete-time linear time invariant (LTI) system, characterized by the state, input, and output matrices F, G, and C,

$$x_{k+1} = Fx_k + Gu_k + \nu_k, \tag{1}$$
$$y_k = Cx_k + \eta_k, \tag{2}$$

with $k \in \mathbb{N}$, the state $x_k \in \mathbb{R}^n$ and output $y_k \in \mathbb{R}^p$ are driven by the control input $u_k \in \mathbb{R}^m$ and i.i.d, Gaussian, and mutually independent process noise $\nu_k \sim \mathcal{N}(0, R_1)$, $R_1 \in \mathbb{S}^n_{++}$, and sensor noise $\eta_k \sim \mathcal{N}(0, R_2)$, $R_2 \in \mathbb{S}^p_{++}$. We assume that F is stable, the pair (F, C) is detectable, and (F, G) is stabilizable.

It is the networked and physically distributed nature of the system that allows the opportunity for the actual measurement y_k to be corrupted by an attack, $\delta_k \in \mathbb{R}^p$. The attack is injected at some point between the measurement and reception of the output by the controller,

$$\bar{y}_k = y_k + \delta_k = Cx_k + \eta_k + \delta_k. \tag{3}$$

If the attacker has access to the measurements, then it is possible for the attack δ_k to cancel some or all of the original measurement y_k.

Because our system is stochastic, we require an estimator to produce a prediction of the system behavior

$$\hat{x}_{k+1} = F\hat{x}_k + Gu_k + L(\bar{y}_k - C\hat{x}_k), \tag{4}$$

where $\hat{x}_k \in \mathbb{R}^n$ is the estimated state and the observer gain L is designed to force the estimate to track the system states.

We consider observer-based feedback controllers

$$u_k = K\hat{x}_k, \tag{5}$$

where $K \in \mathbb{R}^{m\times n}$ is the controller gain matrix. Next, we define the residual sequence

$$r_k = \bar{y}_k - C\hat{x}_k = Ce_k + \eta_k + \delta_k, \tag{6}$$

as the difference between what we actually receive ($\bar{y}_k$) and expect to receive ($C\hat{x}_k$), which evolves according to

$$\begin{aligned} x_{k+1} &= (F + GK)x_k - GKe_k + \nu_k \\ e_{k+1} &= \big(F - LC\big)e_k - L\eta_k + \nu_k - L\delta_k, \end{aligned} \tag{7}$$

where $e_k = x_k - \hat{x}_k$ is the estimation error. In the absence of attacks (i.e., $\delta_k = 0$), we can show that the steady-state distribution of r_k is Gaussian with covariance,

$$\Sigma = \lim_{k\to\infty} \mathbf{E}[r_k r_k^\top] = C\mathbf{P}_e C^\top + R_2, \tag{8}$$

where the steady state covariance of the estimation error $\mathbf{P}_e = \lim_{k\to} \mathbf{E}[e_k e_k^\top]$ is the solution of

$$\mathbf{P}_e = (F - LC)\mathbf{P}_e(F - LC)^\top + LR_2L^\top + R_1. \tag{9}$$

Although similar analysis can be done with other detector choices [5], [6], we consider the chi-squared detector, which constructs a quadratic distance measure z_k and raises alarms when the distance measure exceeds a threshold $\bar{\alpha} \in \mathbb{R}_{>0}$

$$z_k = r_k^\top \Sigma^{-1} r_k > \bar{\alpha} \quad \longrightarrow \quad \text{alarm: } k' = k, \tag{10}$$

such that alarm time(s) k' are produced. The Σ^{-1} factor in the definition of z_k re-scales the distribution ($\mathbf{E}[z_k] = p$, $\mathbf{E}[z_k z_k^\top] = 2p$) so that the threshold $\bar{\alpha}$ can be designed independent of the specific statistics of the noises ν_k and η_k as well as system parameters (e.g., gains K, L) [5].

A. Definition of Attack

Detectors are designed to identify anomalies in system behavior. If an attacker aims to remain undetected, the choice of detector and its parameters limit what the attacker is able to accomplish. The type of attacks we consider here require strong knowledge of and access to system dynamics, statistics of the noises, current estimate ($\hat{x}_k$), and the detector configuration (complete disclosure and model capabilities; complete sensor disruption capabilities). The goal of this powerful stealthy attack is to construct the worst case scenario to aid the design of more robust systems.

Zero-alarm attacks employ attack sequences that maintain the distance measure at or below the threshold of detection, i.e., $z_k \leq \bar{\alpha}$. Hence, these attacks generate no alarms. To satisfy this condition we define the attack as

$$\delta_k = \phi_k - (y_k - C\hat{x}_k) = -Ce_k - \eta_k + \phi_k, \tag{11}$$

where $\phi_k \in \mathbb{R}^p$ is any vector such that $\phi_k^\top \Sigma^{-1} \phi_k \leq \bar{\alpha}$ (recall the attacker has access to the sensor, y_k, and knowledge of the estimator, $\hat{x}_k$). Under this attack no alarms are raised,

$$\begin{aligned} z_k &= r_k^\top \Sigma^{-1} r_k = (Ce_k + \eta_k + \delta_k)^\top \Sigma^{-1}(Ce_k + \eta_k + \delta_k), \\ &= \phi_k^\top \Sigma^{-1} \phi_k \leq \bar{\alpha}. \end{aligned} \tag{12}$$

B. Reachable Set

The dynamics under a zero-alarm attack (11) become

$$\begin{aligned} x_{k+1} &= Fx_k + GK\hat{x}_k + \nu_k, \\ \hat{x}_{k+1} &= LCx_k + (F + GK - LC)\hat{x}_k - LCe_k + L\phi_k, \\ e_{k+1} &= Fe_k - L\phi_k + \nu_k. \end{aligned} \tag{13}$$

We stack these into a combined state $\xi_k = \left[x_k^\top,\, \hat{x}_k^\top,\, e_k^\top\right]^\top$ and combined input $\mu_k = \left[\nu_k^\top,\, \phi_k^\top\right]^\top$,

$$\xi_{k+1} = A\xi_k + B\mu_k, \tag{14}$$

with

$$A = \begin{bmatrix} F & GK & 0 \\ LC & F + GK - LC & -LC \\ 0 & 0 & F \end{bmatrix},\ B = \begin{bmatrix} I & 0 \\ 0 & L \\ I & -L \end{bmatrix}. \tag{15}$$

Remark 1: The choice of including x_k, $\hat{x}_k$, and e_k seems redundant since $e_k = x_k - \hat{x}_k$, however, this choice is crucial as we layer additional constraints into the design optimization. Throughout the rest of the paper we will use a $n\times 3n$ selection matrix $E_x = [I_n,\, 0_n,\, 0_n]$ to pull out the state $x_k = E_x\xi_k$.

The reachable set of attack-induced states is then,

$$\mathcal{R} = \left\{ x_k = E_x \xi_k \;\middle|\; \begin{array}{l} \xi_{\kappa+1} = A\xi_\kappa + B\mu_\kappa, \\ \xi_1 = 0,\ \phi_\kappa^\top \Sigma^{-1} \phi_\kappa \le \bar{\alpha}, \\ \nu_\kappa^\top R_1^{-1} \nu_\kappa \le \bar{\nu},\ \forall k \in \mathbb{N} \end{array} \right\}, \quad (16)$$

where the ellipsoidal bound on the attack ϕ_k is imposed by the attacker's desire to remain stealthy (12). The ellipsoidal bound on the noise is created by truncating the Gaussian process noise to a desired probability, i.e., $\Pr[\nu_k^\top R_1^{-1} \nu_k \le \bar{\nu}] = p_\nu$, where p_ν is some desired (typically high) probability. In principle, the noise has unbounded support, and hence the reachable set is unbounded. To ensure bounded reachable sets, we apply this truncation at the desired confidence level.

III. Scalable Gain Design via LMI Approach

In the first subsection below, we reframe an existing result more concisely, which identifies a minimal outer ellipsoidal bound on the set of states reachable by a stealthy (zero-alarm) attacker. We then consider minimizing the upper bound of this set further through the design of the feedback and estimator gains K and L. A trivial solution exists to this design problem - to make either $GK = 0$ or $L = 0$. Doing so cuts the feedback loop and guarantees that corrupted measurements do not impact the system state. Simultaneously, this destroys the purpose - more specifically the performance - of the feedback loop. While many performance metrics could be used, in Section III-B we impose a $\|H\|_2$ constraint to avoid these trivial solutions. Unlike prior work where the performance criteria ignored the distribution of the noise, this $\|H\|_2$ constraint is specific to the covariance of the noise, thereby allowing our design optimization to leverage this important knowledge. This input covariance based (ICB) $\|H\|_2$ constraint is non-convex; and here we offer a convexification of the ICB $\|H\|_2$ criteria into an LMI framework.

A. Bounding Ellipsoid LMI (given K and L)

Before we move on to the *synthesis* problem of designing the gain matrices, we first provide a solution to the *analysis* problem of finding a minimal outer ellipsoidal bound of the reachable set given K and L, when the system is driven by the process noise and attack. A similar analysis result appears in [3], however, there the problem is split into two optimizations - one to find a bound on the estimation error reachable set, the result of which is used in the second optimization to bound the state reachable set. Here, in Lemma 1, we solve these simultaneously through the stacked states ξ_k and inputs μ_k.

Lemma 1: Given the system matrices A and B in (15), gains K, L, detector threshold $\bar{\alpha}$ with steady state residual covariance Σ, process noise truncation threshold $\bar{\nu}$ with covariance R_1, if there exists a constant $a \in [0, 1)$ such that the following optimization is feasible, the solution of

$$\begin{cases} \min\limits_{a_1, a_2, \mathcal{Q}} \mathbf{tr}(E_x^\top \mathcal{Q} E_x) \\ \quad \text{s.t. } \ 0 \le a_1, a_2 < 1, \quad a_1 + a_2 \ge a, \\ \qquad \mathcal{B} = \begin{bmatrix} a\mathcal{Q} & \mathcal{Q}A^\top & 0 \\ A\mathcal{Q} & \mathcal{Q} & B \\ 0 & B^\top & \frac{1-a}{2-a} W \end{bmatrix} \succeq 0, \end{cases} \quad (17)$$

provides the shape matrix $\mathcal{Q}$ of an ellipsoidal bound on the reachable set of states, i.e., $\mathcal{R} \subseteq \mathcal{E}(E_x^\top \mathcal{Q} E_x)$, where

$$W = \begin{bmatrix} \frac{1-a_1}{\bar{\nu}} R_1^{-1} & 0 \\ 0 & \frac{1-a_2}{\bar{\alpha}} \Pi \end{bmatrix}, \quad \Pi = \Sigma^{-1}. \quad (18)$$

Proof: The stacked dynamics (14) is driven by two inputs which are both ellipsoidally bounded. Letting $W_1 = R_1^{-1}$ and $W_2 = \Sigma^{-1}$, from Lemma 1 in [7], we can write,

$$V_{k+1} - aV_k - \frac{1-b}{\bar{\alpha}} \phi_k^\top \Sigma^{-1} \phi_k - \frac{1-a_1}{\bar{\nu}} \nu_k^\top R_1^{-1} \nu_k \le 0. \quad (19)$$

Substituting the choice $V_k = \xi_k^\top \left(\frac{2-a}{1-a} \mathcal{P}\right) \xi_k \le \frac{2-a}{1-a}$, $\mathcal{P} \succ 0$, into this equation and expanding using the dynamics (14) results in the LMI,

$$\mathcal{H} = \begin{bmatrix} a\mathcal{P} & A^\top \mathcal{P} & 0 \\ \mathcal{P}A & \mathcal{P} & \mathcal{P}B \\ 0 & B^\top \mathcal{P} & \frac{1-a}{2-a} W \end{bmatrix} \succeq 0. \quad (20)$$

which defining $\mathcal{Q} = \mathcal{P}^{-1}$, can be rephrased as $\mathcal{B}$, see Appendix A.

B. Input Covariance Based $\|H\|_2$ Constraint

The introduction of this section and past related work has identified that trivial solutions exist for the synthesis problem unless a performance criteria is imposed in the optimization [7]. One of the distinguishing features of this work is that we implement an input covariance based $\|H\|_2$ constraint, which involves the covariances of the process and sensor noises. In contrast, the standard $\|H\|_2$ metric is independent of the noise distribution and hence is unable to exploit this information to yield improved solutions. The challenge is to convexify and linearize this inherently nonlinear constraint. Most optimizations in the literature either use $\|H\|_\infty$ robust constraint that is already convex [7], [9] or solve the standard $\|H\|_2$ using iterative algorithms [22], [23].

To specify the performance, Robust Control studies the gain observed in a signal $h_k = Hx_k + \theta_k$. Here, for the system *without attack*, we consider the system driven by process and sensor noise and enforce an $\|H\|_2$ constraint between the controlled output h_k with sensor noise $\theta_k \in \mathcal{N}(0, R_3)$, and excitation $\omega_k = \left[\nu_k^\top,\ \eta_k^\top\right]^\top$. When there is no attack the system evolves according to

$$x_{k+1} = Fx_k + GK\hat{x}_k + \nu_k, \quad (21)$$

$$\hat{x}_{k+1} = LCx_k + (F + GK - LC)\hat{x}_k + L\eta_k, \quad (22)$$

$$h_k = Hx_k + \theta_k, \quad (23)$$

which can be combined using the stacked state $\zeta_k = E_{x\hat{x}} \xi_k = \left[x_k^\top,\ \hat{x}_k^\top\right]^\top$, with $E_{x\hat{x}} = [I_{2n},\ 0_{2n \times n}]$,

$$\zeta_{k+1} = \hat{A}\zeta_k + \hat{B}\omega_k, \quad (24)$$

with $\hat{A} = E_{x\hat{x}} A E_{x\hat{x}}^\top$ and $\hat{B} = E_{x\hat{x}} B E_{x\hat{x}}^\top$.

Remark 2: The value of the redundant definition of ξ_k (see Remark 1) is to express the state matrix without attack $\hat{A}$ as a sub-block of the state matrix under attack A. This parallel structure facilitates integrating the $\|H\|_2$ constraint (without attack) with the reachable set calculation (under attack).

The standard $\|H\|_2$ norm is defined as the expected power of the response to zero mean white noise with unit covariance, i.e., $\lim_{k\to\infty} \mathbf{E}[h_k^\top h_k]$ [24]. The ICB $\|H\|_2$ criteria takes the same approach, but instead normalizes by the expected power of the input ω_k and specifies the gain from the noise to the controlled output should be less than a desired value $\bar{\gamma}$,

$$\|H\|_2 = \lim_{k\to\infty} \frac{\mathbf{E}[h_k^\top h_k]}{\mathbf{E}[\omega_k^\top \omega_k]} \leq \bar{\gamma}. \tag{25}$$

Lemma 2: Given the dynamics in (24), the ICB $\|H\|_2$ constraint in (25) is satisfied if the convex inequality holds,

$$\begin{aligned} \mathcal{C}_h = \mathbf{tr}\big(\hat{E}_x^\top H^\top H \hat{E}_x \mathbf{P}\big) + \mathbf{tr}(R_3) \\ - \bar{\gamma}\big(\mathbf{tr}(R_1) + \mathbf{tr}(R_2)\big) \leq 0, \end{aligned} \tag{26}$$

where $\hat{E}_x = [I_n,\, 0_n]$, and,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_x & \mathbf{P}_{x\hat{x}} \\ \mathbf{P}_{x\hat{x}}^\top & \mathbf{P}_{\hat{x}} \end{bmatrix} = \lim_{k\to\infty} \mathbf{P}_k = \lim_{k\to\infty} \mathbf{E}[\zeta_k \zeta_k^\top], \tag{27}$$

is the steady state covariance satisfying the Lyapunov equation

$$\mathbf{P} = \hat{A}\mathbf{P}\hat{A}^\top + \hat{R}, \qquad \mathbf{P} \succeq 0, \quad \hat{R} = \begin{bmatrix} R_1 & 0 \\ 0 & LR_2L^\top \end{bmatrix}. \tag{28}$$

Proof: See Appendix B.

The inequality (26) bounds the actual performance, γ,

$$\gamma = \frac{\mathbf{tr}\left(H\mathbf{P}_x H^\top\right) + \mathbf{tr}(R_3)}{\mathbf{tr}(R_2) + \mathbf{tr}(R_1)} \leq \bar{\gamma}, \tag{29}$$

by the worst allowable performance, $\bar{\gamma}$.

To use this $\|H\|_2$ constraint in a convex optimization we need to linearize the Lyapunov equation constraint. We state this as part of a complete convex optimization problem to design the gains K and L to achieve the optimal $\|H\|_2$ gain.

Theorem 1: Given the convex optimization,

$$\begin{cases} \min\limits_{\mathbf{P}_x, \mathbf{Q}_1, X, Y, Z} \mathbf{tr}(H\mathbf{P}_x H^\top) \\ \text{s.t.} \quad \mathcal{C}_L \succeq 0, \end{cases} \tag{30}$$

$$\mathcal{C}_L = \begin{bmatrix} \mathbf{Q}_1 & I & \mathbf{Q}_1F + XC & Z & \mathbf{Q}_1R_1 & XR_2 \\ * & \mathbf{P}_x & F & F\mathbf{P}_x + GY & R_1 & 0 \\ * & * & \mathbf{Q}_1 & I & 0 & 0 \\ * & * & * & \mathbf{P}_x & 0 & 0 \\ * & * & * & * & R_1 & 0 \\ * & * & * & * & * & R_2 \end{bmatrix}, \tag{31}$$

with solution $\mathbf{P}_x^*$ and given matrices

$$\begin{aligned} &\Gamma_1 = GY(I - \mathbf{Q}_1\mathbf{P}_x)^{-1}, \quad \Gamma_2 = F, \quad \Gamma_4 = -XC, \\ &\Gamma_3 = (\mathbf{Q}_1GY + XC\mathbf{P}_x + \mathbf{Q}_1F\mathbf{P}_x - Z)(I - \mathbf{Q}_1\mathbf{P}_x)^{-1}, \end{aligned} \tag{32}$$

if the nonlinear algebraic Riccati equation,

$$\mathbf{Q}_{12}\Gamma_1\mathbf{Q}_{12} + \mathbf{Q}_{12}\Gamma_2 + \Gamma_3\mathbf{Q}_{12} + \Gamma_4 = 0, \tag{33}$$

has at least one real solution, $\mathbf{Q}_{12}$, then the gains $L = \mathbf{Q}_{12}^{-1}X$, $K = Y(I - \mathbf{Q}_1\mathbf{P}_x)^{-1}\mathbf{Q}_{12}$ result in the optimal $\|H\|_2$ gain,

$$\gamma^* = \frac{\mathbf{tr}\big(H\mathbf{P}_x^* H^\top\big) + \mathbf{tr}(R_3)}{\mathbf{tr}(R_2) + \mathbf{tr}(R_1)}. \tag{34}$$

Proof: See Appendix C

Remark 3: If the NARE (33) does not provide a real solution, there are several existing scalable nonlinear iterative algorithms (e.g., [22], [9]) that can be employed to approximate γ^* and the corresponding gains.

Remark 4: The NARE (33) can have more than one real solution (the Schur Algorithm in [25] computes all $\binom{2n}{n}$ possible solutions) and each corresponds to different gains K and L that result in the same optimal γ^*. Out of these solutions, we select the one that leads to the smallest attacked reachable set (using Lemma 1). In practice we have observed far fewer - namely two - real solutions. Appendix D presents an analysis of this particular scenario.

The optimal performance γ^* corresponds to gains that prioritize performance over security. At the other end of the performance spectrum, security increases (reachable set size decreases) as either L or GK approach zero. This represents worst-case optimal performance (without closed loop control) and Lemma 3 provides the corresponding value of $\gamma = \gamma_0$.

Lemma 3: Given the system dynamics (21) and (22), the open loop ICB $\|H\|_2$ gain γ_0 is given by

$$\gamma_0 = \frac{\mathbf{tr}\big(H\mathbf{P}_x H^\top\big) + \mathbf{tr}(R_3)}{\mathbf{tr}(R_2) + \mathbf{tr}(R_1)}, \tag{35}$$

where the steady state covariance, $\mathbf{P}_x$, is the solution of

$$F\mathbf{P}_x F^\top - \mathbf{P}_x + R_1 = 0. \tag{36}$$

Proof. See Appendix E.

If we pick a performance bound $\bar{\gamma} > \gamma_0$, the goal to maximize security while keeping $\gamma \leq \bar{\gamma}$ will still yield $\gamma = \gamma_0$, since security cannot be improved further once the feedback loop is disconnected. Therefore, for all choices of $\bar{\gamma} \in [\gamma^*,\ \infty]$, the solution for performance $\gamma < \bar{\gamma}$ always lies within the *trade-off interval*, $\gamma^* \leq \gamma \leq \gamma_0$.

C. Bounding ellipsoid LMI (designing K and L)

The goal of this paper is to construct an optimization to design K and L such that the impact of an attacker on the proposed upper bound over the the reachable states is minimized. However, when K and L are considered variables of the Lemma 1 optimization, (17) contains nonlinear terms. In the sections that follow, we impose some structure on the solution so that we can linearize the overall design problem. Each choice will be motivated individually, but it is also the combined effect of the these structures taken together that yield the final *linear* matrix inequality. While they do impose some structure on the solution to accomplish the linearization, they are best seen as one choice out of typically many quasi-optimal solutions which still enforce the original nonlinear constraints.

Imposed Structure 1: There are an infinite number of differently shaped tight (tangent) outer ellipsoidal bounds of the stacked reachable states ξ. Of these we select one that satisfies the following structure for the inverse of the shape matrix, which assumes the independence of the ellipsoidal bound on the estimation error e_k from the ellipsoidal bound on the combined state x_k and estimate $\hat{x}_k$,

$$\mathcal{Q}^{-1} = \mathcal{P} = \begin{bmatrix} \mathcal{P}_1 & \mathcal{P}_{12} & \\ \mathcal{P}_{12}^\top & \mathcal{P}_2 & \\ & & \mathcal{P}_3 \end{bmatrix}. \tag{37}$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3229774

This is inspired by a similar assumption made in [7]. This selection enables us to utilize the parallel dynamics with and without attack (see Remarks 1 and 2) and linearize the original LMI with respect to K and L. This selection also permits inverting each block separately, such that $\mathcal{P}_3^{-1} = \mathcal{Q}_e$ and

$$\begin{bmatrix} \mathcal{P}_1 & \mathcal{P}_{12} \\ \mathcal{P}_{12}^\top & \mathcal{P}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathcal{Q}_x & \mathcal{Q}_{x\hat{x}} \\ \mathcal{Q}_{x\hat{x}}^\top & \mathcal{Q}_{\hat{x}} \end{bmatrix}. \tag{38}$$

With the linearizing change of coordinates used in [10], [7],

$$T_2 = \begin{bmatrix} T_3 & & \\ & T_3 & \\ & & I_n \end{bmatrix}, \qquad T_3 = \begin{bmatrix} \mathcal{Q}_x & I & 0 \\ \mathcal{Q}_{x\hat{x}}^\top & 0 & 0 \\ 0 & 0 & I \end{bmatrix}, \tag{39}$$

the LMI $\mathcal{H}$, (20), can be transformed as

$$\mathcal{H}_L = T_2^\top \mathcal{H} T_2 = \begin{bmatrix} a\mathcal{P}_L & A_L^\top & 0 \\ A_L & \mathcal{P}_L & B_L \\ 0 & B_L^\top & \frac{1-a}{2-a} W \end{bmatrix}, \tag{40}$$

with $\mathcal{P}_L = T_3^\top \mathcal{P} T_3$, $B_L = T_3^\top \mathcal{P} B$, and $A_L = T_3^\top \mathcal{P} A T_3$, where $Y_1 = \mathcal{P}_{12} L$, $X_1 = K \mathcal{Q}_{x\hat{x}}^\top$,

$$\begin{aligned} \mathcal{P}_L &= \begin{bmatrix} \mathcal{Q}_x & I & 0 \\ I & \mathcal{P}_1 & 0 \\ 0 & 0 & \mathcal{P}_3 \end{bmatrix}, \; B_L = \begin{bmatrix} I & 0 \\ \mathcal{P}_1 & Y_1 \\ \mathcal{P}_3 & -\mathcal{P}_3 L \end{bmatrix}, \\ A_L &= \begin{bmatrix} F\mathcal{Q}_x + GX_1 & F & 0 \\ Z_1 & \mathcal{P}_1 F + Y_1 C & -Y_1 C \\ 0 & 0 & \mathcal{P}_3 F \end{bmatrix}, \\ Z_1 &= \mathcal{P}_1 F \mathcal{Q}_x + \mathcal{P}_{12} L C \mathcal{Q}_x + \mathcal{P}_1 G K \mathcal{Q}_{x\hat{x}}^\top + \mathcal{P}_{12} F \mathcal{Q}_{x\hat{x}}^\top \\ &\quad + \mathcal{P}_{12} G K \mathcal{Q}_{x\hat{x}}^\top - \mathcal{P}_{12} L C \mathcal{Q}_{x\hat{x}}^\top. \end{aligned} \tag{41}$$

One of the useful features of this transformation is that $\mathcal{Q}_x = E_x^\top \mathcal{Q} E_x$, the quantity used in the objective function of Lemma 1, appears as a variable of the LMI. This section provides the linearization necessary to separate the gains K and L as variables in Lemma 1 (and could then be used as the starting point if a different performance criteria was used, as opposed to the $\|H\|_2$ constraint considered in this paper).

Note that the stability of the closed loop system is implicitly guaranteed if $\mathcal{H}_L > 0$ and $\mathcal{C}_L > 0$ (see Appendix F).

D. Combining Performance and Security

We design the controller and estimator gains to minimize the impact of attacks on the system state, measured by an outer ellipsoidal bound on the reachable states when the system is driven by the attack and process noise. There are an infinite number of potential outer bounding - and tight - ellipsoids on the reachable set, therefore to combine the LMI constraints from the reachable set and $\|H\|_2$ calculations, we make a specific choice about the outer ellipsoidal bound we select.

Imposed Structure 2: We select the shape matrix of the ellipsoidal bound of the states x_k and estimate $\hat{x}_k$ under attack $E_{x\hat{x}}^\top \mathcal{Q} E_{x\hat{x}}$ - see (38) - to have the same orientation as the covariance of the states and estimate without attack (ζ_k), and the shape matrix of the estimation error e_k under attack to have the same orientation of the estimation error without attack,

$$\sigma_1 \mathbf{P} = E_{x\hat{x}}^\top \mathcal{Q} E_{x\hat{x}}, \qquad \sigma_2 \mathbf{P}_e = E_e^\top \mathcal{Q} E_e \tag{42}$$

where $\sigma = [\sigma_1, \; \sigma_2]^\top$ is the vector of scaling factors that become new variables of the method and is a function of gains $(L, \; K)$. Since $\mathbf{Q} = \mathbf{P}^{-1}$, this sets up a common set of variables to link the $\|H\|_2$ (left) and ellipsoidal bound (right) constraints,

$$\begin{aligned} &\sigma_1 \begin{bmatrix} \mathbf{P}_x & \mathbf{P}_{x\hat{x}} \\ \mathbf{P}_{x\hat{x}}^\top & \mathbf{P}_{\hat{x}} \end{bmatrix} = \begin{bmatrix} \mathcal{Q}_x & \mathcal{Q}_{x\hat{x}} \\ \mathcal{Q}_{x\hat{x}}^\top & \mathcal{Q}_{\hat{x}} \end{bmatrix}, \quad \sigma_2 \mathbf{P}_e = \mathcal{Q}_e \\ &\underbrace{\frac{1}{\sigma_1} \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \mathbf{Q}_2 \end{bmatrix}}_{\|H\|_2} = \underbrace{\begin{bmatrix} \mathcal{P}_1 & \mathcal{P}_{12} \\ \mathcal{P}_{12}^\top & \mathcal{P}_2 \end{bmatrix}}_{\mathcal{E}(\mathcal{Q})}, \quad \frac{1}{\sigma_2} \mathbf{P}_e^{-1} = \mathcal{P}_3. \end{aligned} \tag{43}$$

The structure above allows us to replace variables in the ellipsoidal bound optimization $\mathcal{Q}$ and $\mathcal{P}$ with quantities from the performance criteria, $\mathbf{P}$ and $\mathbf{Q}$, respectively.

Based on (43) we can link the variables of the bounding ellipsoid LMI with the $\|H\|_2$ constraint,

$$\begin{aligned} X &= \sigma_1 Y_1 = \mathbf{Q}_{12} L, \quad Y = \frac{X_1}{\sigma_1} = K \mathbf{P}_{x\hat{x}}^\top, \\ Z &= Z_1 = \mathbf{Q}_1 F \mathbf{P}_x + XC\mathbf{P}_x + \mathbf{Q}_1 GY + \mathbf{Q}_{12} F \mathbf{P}_{x\hat{x}}^\top \\ &\quad + \mathbf{Q}_{12} GY - XC \mathbf{P}_{x\hat{x}}^\top. \end{aligned} \tag{44}$$

Defining the new change of variables,

$$\mathbf{P}_{x\sigma} = \sigma_1 \mathbf{P}_x, \; \mathbf{Q}_{1\sigma} = \tfrac{1}{\sigma_1} \mathbf{Q}_1, \; \mathbf{X}_\sigma = \tfrac{1}{\sigma_1} \mathbf{X}, \; \mathbf{Y}_\sigma = \sigma_1 \mathbf{Y}, \tag{45}$$

we can rewrite A_L, B_L , $\mathcal{P}_L$ as

$$\begin{aligned} A_L &= \begin{bmatrix} F\mathbf{P}_{x\sigma} + GY_\sigma & F & 0 \\ Z & \mathbf{Q}_{1\sigma} F + X_\sigma C & -X_\sigma C \\ 0 & 0 & \mathcal{P}_3 F \end{bmatrix}, \\ B_L &= \begin{bmatrix} I & 0 \\ \mathbf{Q}_{1\sigma} & X_\sigma \\ \mathcal{P}_3 & -\mathcal{P}_3 L \end{bmatrix}, \quad \mathcal{P}_L = \begin{bmatrix} \mathbf{P}_{x\sigma} & I & 0 \\ I & \mathbf{Q}_{1\sigma} & 0 \\ 0 & 0 & \mathcal{P}_3 \end{bmatrix}. \end{aligned} \tag{46}$$

In order to make the convex constraints $\mathcal{C}_L$ and $\mathcal{C}_h$ compatible with this new change of variables we apply the transformations, $\mathcal{C}_{L\sigma} = T_4^\top \mathcal{C}_L T_4$ and $\mathcal{C}_{h\sigma} = \sigma_1 \mathcal{C}_h$, with $T_4 = \sqrt{\sigma_1}\mathbf{diag}\big(\big[\frac{1}{\sigma_1}, 1, \frac{1}{\sigma_1}, 1, 1, 1\big]\big)$,

$$\mathcal{C}_{L\sigma} = \begin{bmatrix} \mathbf{Q}_{1\sigma} & I & \mathbf{Q}_{1\sigma} F + X_\sigma C & Z & \sigma_1 \mathbf{Q}_{1\sigma} R_1 & \sigma_1 X_\sigma R_2 \\ * & \mathbf{P}_{x\sigma} & F & F\mathbf{P}_{x\sigma} + GY_\sigma & \sigma_1 R_1 & 0 \\ * & * & \mathbf{Q}_{1\sigma} & I & 0 & 0 \\ * & * & * & \mathbf{P}_{x\sigma} & 0 & 0 \\ * & * & * & * & \sigma_1 R_1 & 0 \\ * & * & * & * & * & \sigma_1 R_2 \end{bmatrix}.$$

$$\mathcal{C}_{h\sigma} = \mathbf{tr}\big(H \mathbf{P}_{x\sigma} H^\top\big) + \sigma_1 \mathbf{tr}(R_3) - \sigma_1 \bar{\gamma}\big(\mathbf{tr}(R_1) + \mathbf{tr}(R_2)\big)$$

The approach offered in [26], relies on an iterative scheme to avoid the nonlinearities surrounding the observer gain matrix L. In each iteration, the solution of the proposed convex optimization is used to find L using the Riccati equation in (33) and the corresponding residual covariance Σ using the Lyapunov equation in (9). To eliminate the iterative approach, both of these nonlinearities must be linearized. Another important nonlinearity is the $-\mathcal{P}_3 L$ term within $\mathcal{H}_L$, specifically B_L in (46). We accomplish this linearization through the selection of additional structure, which we will show only marginally reduces the quality of the solutions found.

Imposed Structure 3: We impose $\mathbf{P}_{x\hat{x}} = \mathbf{P}_{\hat{x}}$. This implies $\mathbf{P}_e = \mathbf{P}_x - \mathbf{P}_{\hat{x}} \succeq 0$ and hence $\mathbf{P}_x \succeq \mathbf{P}_{\hat{x}} \succ 0$. This choice, in combination with the objective function $\mathbf{tr}(\mathbf{P}_x)$, serves to

enforce the quality of the estimator (through the choice of the gain L) by reducing the covariance of the estimation error in the absence of attack, $\mathbf{P}_e$, by forcing $\mathbf{P}_x$ to be as close as possible to $\mathbf{P}_{\hat{x}}$.

From Imposed Structure 3 and the identity $\mathbf{PQ} = I$ (see Appendix G) we can show,

$$\mathbf{P}_e^{-1} = \mathbf{Q}_1 = -\mathbf{Q}_{12}, \tag{47}$$

where the second part provides a linear relationship between $\mathbf{Q}_1$ and $\mathbf{Q}_{12}$ which is primarily what linearizes the NARE in (33), and the first part which is a relation between $\mathbf{P}_e$ and $\mathbf{Q}_1$, in companion with Imposed Structure 2, enables us to replace the nonlinear term $-\mathcal{P}_3 L$ with the linear term $\frac{1}{\sigma_2}X$. Therefore, if we define $\tau = \sigma_1/\sigma_2$ we can rewrite matrices A_L, B_L, $\mathcal{P}_L$ as

$$\begin{aligned} A_L &= \begin{bmatrix} F\mathbf{P}_{x\sigma} + GY_\sigma & F & 0 \\ Z & \mathbf{Q}_{1\sigma}F + X_\sigma C & -X_\sigma C \\ 0 & 0 & \tau\mathbf{Q}_{1\sigma}F \end{bmatrix}, \\ B_L &= \begin{bmatrix} I & 0 \\ \mathbf{Q}_{1\sigma} & X_\sigma \\ \tau\mathbf{Q}_{1\sigma} & \tau X_\sigma \end{bmatrix}, \ \mathcal{P}_L = \begin{bmatrix} \mathbf{P}_{x\sigma} & I & 0 \\ I & \mathbf{Q}_{1\sigma} & 0 \\ 0 & 0 & \tau\mathbf{Q}_{1\sigma} \end{bmatrix}. \end{aligned} \tag{48}$$

The relation in (47) between $\mathbf{P}_e$ and $\mathbf{Q}_1$ also helps to integrate the Lyapunov equation in (9) into the convex optimization. Lemma 4 relates the matrix $\Pi = \Sigma^{-1}$ with the existing decision variables.

Lemma 4: The positive definite matrix Π, can be expressed as $\Pi = \Sigma^{-1}$ where Σ comes from (8) and (9), if,

$$\mathcal{X}_{L\sigma} = \begin{bmatrix} \Pi & \Pi C & \Pi R_2 \\ C^\top\Pi & \sigma_1\mathbf{Q}_{1\sigma} & 0 \\ R_2\Pi & 0 & R_2 \end{bmatrix} \succeq 0 \quad \text{and} \tag{49}$$

$$\mathcal{S}_{L\sigma} = \begin{bmatrix} \mathbf{Q}_{1\sigma} & \mathbf{Q}_{1\sigma}F + X_\sigma C & \sigma_1\mathbf{Q}_{1\sigma}R_1 & \sigma_1 X_\sigma R_2 \\ (\mathbf{Q}_{1\sigma}F + X_\sigma C)^\top & \mathbf{Q}_{1\sigma} & 0 & 0 \\ \sigma_1 R_1\mathbf{Q}_{1\sigma} & 0 & \sigma_1 R_1 & 0 \\ \sigma_1 R_2 X_\sigma^\top & 0 & 0 & \sigma_1 R_2 \end{bmatrix} \succeq 0. \tag{50}$$

Proof: See Appendix H.

Remark 5: Both this work and [7] employ a detector that forms a quadratic distance measure of the form $z_k = r_k^\top R^{-1} r_k$. In [7], the authors make the assumption that the normalizing factor R is the shape matrix of an ellipsoid that bounds the residual, but is not related to the actual covariance and does not offer guarantees about tightness. This makes the gain design problem easier, however, it means that the false alarm rate of the detector would change as the gains K and L are designed. In this paper, we stay true to the definition of the chi-squared detector where $R = \Sigma$, the true covariance of the steady state residual. Doing so, however, introduces nonlinearities in the design optimization since Σ depends on the observer gain L through (8)-(9). Therefore, it is a critical feature of our formulation that it approximates Σ (the true residual covariance) through the LMIs in Lemma 4. This approach ensures that the distance measure distribution remains chi-squared and the false alarm rate is constant. It is this separation between the detector tuning and the gain optimization that enables showing in Section IV that the optimization is entirely independent from the threshold selection and false alarm rate.

Algorithm 1: $K, L = \text{Thm } 2(F, G, C, H, R_1, R_2, \epsilon)$

$\tau \leftarrow 1$
$\sigma_1 \leftarrow$ initialize σ_1 such that (51) is feasible.
while *true* **do**
  $\mathbf{P}_{x\sigma}, \mathbf{Q}_{1\sigma}, X_\sigma, Y_\sigma, Z \leftarrow$ fix (σ_1, τ); solve (51)
  $(\hat{\sigma}_1, \hat{\tau}) \leftarrow$ fix $(\mathbf{Q}_{1\sigma}, X_\sigma)$; solve (51)
  if $|\sigma_1 - \hat{\sigma}_1| < \epsilon$ **then**
    $\mathbf{P}_x, \mathbf{Q}_1, X, Y, Z \leftarrow$ (45)
    return $L \leftarrow -\mathbf{Q}_1^{-1}X, \ K \leftarrow Y\left(\mathbf{P}_x - \mathbf{Q}_1^{-1}\right)^{-1}$
  else $(\sigma_1, \tau) \leftarrow (\hat{\sigma}_1, \hat{\tau})$
end

Combining these linearizations, we provide Theorem 2.

Theorem 2: Consider a LTI system (1) equipped with a chi-squared detector (10). If there exists constants $a, a_2 \in [0, 1)$, that make the following optimization feasible then the solution of Algorithm 1 over the convex optimization

$$\begin{cases} \min\limits_{\substack{a_1, \mathbf{P}_{x\sigma}, \mathbf{Q}_{1\sigma}, \Pi \\ X_\sigma, Y_\sigma, Z}} \mathbf{tr}(\mathbf{P}_{x\sigma}) \\ \quad \text{s.t.} \quad 0 \le a_1 < 1, \quad a_1 + a_2 \ge a, \\ \qquad \mathcal{H}_L, \mathcal{C}_{L\sigma}, \mathcal{S}_{L\sigma}, \mathcal{X}_{L\sigma} \succeq 0, \ \mathcal{C}_{h\sigma} \le 0, \end{cases} \tag{51}$$

provides the observer L and controller K gain matrices that minimize the outer ellipsoidal bound (subject to Imposed Structures 1, 2, and 3) on the set of states reachable by a zero-alarm attack (12) while maintaining an ICB $\|H\|_2$ gain (25) no bigger than $\bar{\gamma} \in [\gamma^*, \gamma_0]$.

Proof: See Appendix I

IV. Gain Design via Geometric Approach

We now develop an alternative nonlinear optimization approach to the gain design problem based on ellipsoidal geometric tools. Although less scalable, the high accuracy of the method is able to characterize the trade-off between robustness and security with greater fidelity. It is the true nature of this trade-off curve that demonstrates giving up a small amount of $\|H\|_2$ performance can lead to large gains in our reachability-based security metric.

In the steady state condition (once the transience due to the initial condition is gone), the solution for state and estimation error in (13), for $k \ge 2$, is,

$$e_k = \sum_{i=1}^{k-1} F^{i-1}\nu_{k-i} - F^{i-1}L\phi_{k-i}, \tag{52}$$

$$x_k = \sum_{i=1}^{k-1} F^{i-1}\nu_{k-i} + \left((F+GK)^{i-1} - F^{i-1}\right)L\phi_{k-i}. \tag{53}$$

Solving for x_k in (53) allows us to compute the reachable set at time k as the geometric sum of the history of ellipsoidally-bounded inputs ν_k (noise) and ϕ_k (attack) [27],

$$\mathcal{R}_k = \bigoplus_{i=0}^{k-2} \mathcal{E}\left(\bar{\nu}F^i R_1 F^{i\top}\right) \oplus \mathcal{E}\left(\bar{\alpha}H_i L\Sigma L^\top H_i^\top\right), \tag{54}$$

$$\mathbf{P}_{g,L}^{ij} = \sum_{q=0}^{k^*} A_g^q \begin{bmatrix} 0 & 0 \\ 0 & J_L^{ij} R_2 L^\top + L R_2 (J_L^{ij})^\top \end{bmatrix} A_g^{q\top} - \sum_{q=1}^{k^*} \left(\sum_{r=1}^{q} A_g^{r-1} \begin{bmatrix} 0 & 0 \\ 0 & J_L^{ij} C \end{bmatrix} A_g^{q-r} R A_g^{q\top} + \sum_{r=1}^{q} \left(A_g^{r-1} \begin{bmatrix} 0 & 0 \\ 0 & J_L^{ij} C \end{bmatrix} A_g^{q-r} R A_g^{q\top} \right)^\top \right)$$

$$\mathbf{P}_{g,K}^{uv} = \sum_{q=1}^{k^*} \left(\sum_{r=1}^{q} A_g^{r-1} \begin{bmatrix} G J_K^{uv} & -G J_K^{uv} \\ 0 & 0 \end{bmatrix} A_g^{q-r} R A_g^{q\top} + \sum_{r=1}^{q} \left(A_g^{r-1} \begin{bmatrix} G J_K^{uv} & -G J_K^{uv} \\ 0 & 0 \end{bmatrix} A_g^{q-r} R A_g^{q\top} \right)^\top \right), \quad \begin{matrix} i \in \overline{1,n},\ j \in \overline{1,p} \\ u \in \overline{1,m},\ v \in \overline{1,n} \end{matrix} \quad (\star)$$

$$\sum_{q=1}^{k^*} \frac{\mathbf{tr}\left((H_q^\top H_q)(J_L^{ij} \Sigma L^\top + L C E_e \mathbf{P}_{g,L}^{ij} E_e^\top C^\top L^\top + L \Sigma J_L^{ij T}) \right)}{2\sqrt{\mathbf{tr}(H_q L \Sigma L^\top H_q^\top)}} + \lambda\, \mathbf{tr}(H \hat{E}_x \mathbf{P}_{g,L}^{ij} \hat{E}_x^\top H^\top) = 0, \quad \begin{matrix} i \in \overline{1,n},\ j \in \overline{1,p} \\ E_e = [0_n,\ I_n] \end{matrix} \quad (\dagger)$$

$$\sum_{q=1}^{k^*} \frac{\sum_{r=1}^{q} \mathbf{tr}\left(2 L \Sigma L^\top H_q^\top (F+GK)^{r-1} (G J_K^{uv}) (F+GK)^{q-r} \right)}{2\sqrt{\mathbf{tr}(H_q L \Sigma L^\top H_q^\top)}} + \lambda\, \mathbf{tr}(H \hat{E}_x \mathbf{P}_{g,K}^{uv} \hat{E}_x^\top H^\top) = 0, \quad \begin{matrix} u \in \overline{1,m},\ v \in \overline{1,n} \end{matrix} \quad (\ddagger)$$

with $H_i = (F+GK)^i - F^i$, and $\oplus$ denotes geometric sum and $\bigoplus$ denotes geometric series.

The reachable set (54) is a convex, although typically complex, set. Approximations of outer ellipsoidal bounds on this complex shape can be computed iteratively [28] or directly [29] using the set of shape matrices

$$\mathcal{A} := \left\{ \bar{\nu} F^i R_1 F^{i\top},\ \bar{\alpha} H_i L \Sigma L^\top H_i^\top \right\}_{i=0}^{k^*}. \quad (55)$$

Here we use the outer ellipsoidal bound that has the minimum trace (of the shape matrix) such that $\mathcal{R}_k \subseteq \mathcal{E}(Q^*)$, with

$$Q^* = \left(\sum_{i=0}^{k} \sqrt{\bar{\alpha} \mathbf{tr}\left(H_i L \Sigma L^\top H_i^\top \right)} + \sqrt{\bar{\nu} \mathbf{tr}\left(F^i R_1 F^{i\top} \right)} \right) \times \left(\sum_{i=0}^{k} \frac{\bar{\alpha} H_i L \Sigma L^\top H_i^\top}{\sqrt{\mathbf{tr}\left(\bar{\alpha} H_i L \Sigma L^\top H_i^\top \right)}} + \frac{\bar{\nu} F^i R_1 F^{i\top}}{\sqrt{\mathbf{tr}\left(\bar{\nu} F^i R_1 F^{i\top} \right)}} \right). \quad (56)$$

To capture the infinite horizon reachable set, we typically take $k \geq k^*$, where k^* is the settling time of the system such that the transformed noise and attack ellipsoids (55) for $\overline{0, k-k^*}$ are negligible if the system is stable[1] [27].

To design the gain matrices to minimize the reachable set by minimizing the ellipsoid $\mathcal{E}(Q^*)$, we formulate a nonlinear optimization from closed-form derivatives of the ICB $\|H\|_2$ constraint and the minimum trace objective function.

A. ICB $\|H\|_2$ Constraint

We now revisit the $\|H\|_2$ constraint to formulate it for the nonlinear optimization problem of the geometric approach. The attack-free dynamics of the system are

$$\xi_{k+1} = A_g \xi_k + B_g \omega_k, \quad (57)$$

when expressed in terms of $\xi_k = \left[x_k^\top,\ e_k^\top \right]^\top$, with

$$A_g = \begin{bmatrix} F+GK & -GK \\ 0 & F-LC \end{bmatrix}, \quad B_g = \begin{bmatrix} I & 0 \\ I & -L \end{bmatrix}. \quad (58)$$

The steady state covariance of ξ is $\mathbf{P}_g \succeq 0$,

$$\mathbf{P}_g = \begin{bmatrix} \mathbf{P}_x & \mathbf{P}_{xe} \\ \mathbf{P}_{xe}^\top & \mathbf{P}_e \end{bmatrix} = \lim_{k \to \infty} \mathbf{E}[\xi_k \xi_k^\top], \quad (59)$$

and satisfies the Lyapunov equation

$$\mathbf{P}_g = A_g \mathbf{P}_g A_g^\top + R, \quad R = \begin{bmatrix} R_1 & R_1 \\ R_1 & R_1 + L R_2 L^\top \end{bmatrix}. \quad (60)$$

[1]The reachable set is not necessarily contained within the truncated bound, but the violation can be made arbitrarily small with increasing k^*.

Following the same logic that leads to (26) (see Appendix B) but now for $\mathbf{P}_g$, the following inequality,

$$\mathcal{C}_g = \mathbf{tr}\left(\hat{E}_x^\top H^\top H \hat{E}_x \mathbf{P}_g \right) + \mathbf{tr}(R_3) - \bar{\gamma}\left(\mathbf{tr}(R_1) + \mathbf{tr}(R_2) \right) \leq 0, \quad (61)$$

is also valid, because $\hat{E}_x^\top \mathbf{P} \hat{E}_x = \hat{E}_x^\top \mathbf{P}_g \hat{E}_x$. The LMI approach of Theorem 1 provides a solution for $\nabla \mathbf{tr}(H \mathbf{P}_x H^\top) = 0$, yielding the optimal state covariance $\mathbf{P}_x$ and corresponding control gains L and K. This gradient equation can also be expressed in a polynomial form as,

$$\begin{cases} \mathbf{tr}(H \hat{E}_x \mathbf{P}_{g,L}^{ij} \hat{E}_x^\top H^\top) = 0, & i \in \overline{1,n},\ j \in \overline{1,p}, \\ \mathbf{tr}(H \hat{E}_x \mathbf{P}_{g,K}^{uv} \hat{E}_x^\top H^\top) = 0, & u \in \overline{1,m},\ v \in \overline{1,n}. \end{cases} \quad (62)$$

where the optimal gains L and K, are also the solution of $np+mn$ nonlinear equations, where $\mathbf{P}_{g,L}^{ij}$ and $\mathbf{P}_{g,K}^{uv}$ are partial derivatives of $\mathbf{P}_g$ with respect to L, K, given in $(\star)$.

B. Design

We now develop the nonlinear optimization to find the optimal gains L and K to minimize the proposed ellipsoidal bound over the reachable set (56), while maintaining the ICB $\|H\|_2$ gain less than a desired threshold $\bar{\gamma}$. As before, we minimize the trace of the ellipsoid that bounds the reachable set, $\mathbf{tr}(Q^*)$, as a proxy for the impact of the attack, i.e., we aim to minimize

$$\sqrt{\mathbf{tr}(Q^*)} = \sum_{i=0}^{k^*} \left(\sqrt{\bar{\alpha}} \sqrt{\mathbf{tr}\left(H_i L \Sigma L^\top H_i^\top \right)} + \sqrt{\bar{\nu} \mathbf{tr}\left(F^i R_1 F^{i\top} \right)} \right), \quad (63)$$

which can be computed by taking $\mathbf{tr}(.)$ from both sides of (56), and decomposes the ellipsoidal bound on the reachable set into contributions from the noise and attack.

Remark 6: Because the portion of the impact due to noise in (63) (second term) is not a function of gain matrices L and K, it does not play a role in minimizing the attack impact. The threshold $\bar{\alpha}$, which is selected based on the noise distribution and the detector's desired false alarm rate, is also not a function of L and K and since it appears as a uniform scaling factor it also does not play a role in minimizing the attack impact. These are both distinct advantages for selecting an objective that minimizes the *trace* of the shape matrix of the ellipsoidal bound (as opposed to, e.g., minimizing the volume).

From Remark 6, we can reduce (63) to the following objective function

$$\mathcal{J} = \sum_{i=0}^{k^*} \sqrt{\mathbf{tr}\left(H_i L \Sigma L^\top H_i^\top \right)}. \quad (64)$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3229774

Another consequence of using the trace is that the noise truncation probability p_ν, which sets the confidence level of the reachable set, also does not appear in $\mathcal{J}$. This shows that the optimal gains can be computed independent of p_ν, and thus are the optimal gain matrices for any truncation level.

It is intuitive, and can be seen in (64), that over the trade-off interval $\bar{\gamma} \in [\gamma^*, \gamma_0]$ the objective $\mathcal{J}$ decreases as, e.g., L approaches zero, however, the ICB $\|H\|_2$ gain γ is increasing as it approaches γ_0. Thus we expect for $\bar{\gamma} \in [\gamma^*, \gamma_0]$ that the optimal L^* and K^* to minimize the attack impact $\mathcal{J}$ will occur at $\gamma = \bar{\gamma}$. We use this observation to change (61) from an inequality to equality $\mathcal{C}_g = 0$, which simplifies the process of including the ICB $\|H\|_2$ constraint as a Lagrange multiplier in the optimization.

Theorem 3: Consider a LTI system (1) equipped with a chi-squared detector (10). The solution of the optimization

$$\min_{L,K,\lambda} \ \mathcal{J} + \lambda \mathcal{C}_g, \tag{65}$$

can be found by the simultaneous solution of (†), (‡) and

$$\mathbf{tr}\big(H\hat{E}_x \mathbf{P}_g \hat{E}_x^\top H^\top\big) + \mathbf{tr}(R_3) - \bar{\gamma}\big(\mathbf{tr}(R_1) + \mathbf{tr}(R_2)\big) = 0, \tag{66}$$

and provides the observer L and controller K gain matrices that minimize the outer ellipsoidal bound on the set of states reachable by a zero-alarm attack (12) while maintaining an ICB $\|H\|_2$ gain (25) no bigger than $\bar{\gamma} \in [\gamma^*, \gamma_0]$.

Proof: Defining the combined objective as $\Omega = \mathcal{J} + \lambda \mathcal{C}_g$, the necessary condition for a local minimum is that the values of K, L, and λ satisfy $\nabla\Omega = \begin{pmatrix} \frac{\partial \Omega}{\partial L_{ij}} & \frac{\partial \Omega}{\partial K_{uv}} & \frac{\partial \Omega}{\partial \lambda} \end{pmatrix} = 0$, where the equations in (†), (‡), and (66) correspond to each term of the gradient being zero, respectively.

Note that the stabilizability and detectability of the system along with the output covariance being bounded (ensured by the $\|H\|_2$ constraint) guarantees the state covariance is bounded - hence the optimal L and K found through this theorem ensure a stable closed loop system. ■

V. Numerical Example

We present the solutions provided by our methods for two numerical examples. The first system enables us to demonstrate a comparison between the two approaches presented here and the iterative approach from [26] in their ability to quantify the trade-off between performance (ICB $\|H\|_2$ gain) and security (trace of the outer ellipsoid bound on the reachable set). The second system helps to demonstrate the scalability of the fully convexified optimization. Throughout, we employ YALMIP for semi-definite programming equipped with the MOSEK solver [30], [31]. The code to reproduce these results is located online: http://justinruths.com/tcns-codesign/.

A. Performance-Security Trade-off

We consider an LTI system with matrices,

$$F = \begin{bmatrix} 1.04 & -0.14 \\ 0.30 & 0.63 \end{bmatrix}, \ G = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}, \ C = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}, \tag{67}$$
$$R_1 = \begin{bmatrix} 0.018 & -0.022 \\ -0.022 & 0.026 \end{bmatrix}, \ R_2 = \begin{bmatrix} 0.0018 & 0.0031 \\ 0.0031 & 0.0096 \end{bmatrix},$$

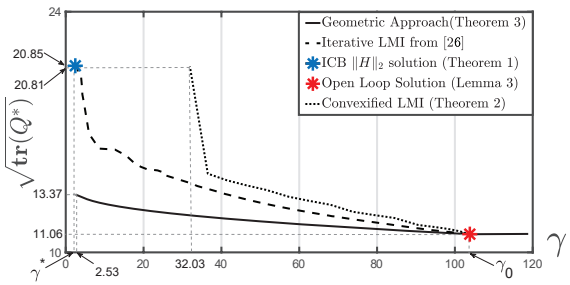

Figure 1: Performance-security curve for the nonlinear geometric approach (solid), fully convex LMI approach (dotted), and the iterative LMI scheme from [26]. The geometric solutions provide the true structure of the fundamental trade-off between performance (x-axis) and security (y-axis). The steps taken to linearize the problem make the convex optimization feasible for only a portion of the entire trade-off interval.

equipped with a chi-squared detector tuned to a false alarm rate of 1% ($\bar{\alpha} = 9.21$), and with process noise truncated at $p_\nu = 99\%$ ($\bar{\nu} = 9.21$). We select a settling time of $k^* = 35$. In addition we select the controlled measurement $h_k = y_k$. We start by computing the optimal ICB $\|H\|_2$ gain $\gamma^* = 2.45$ using Theorem 1. The solution for gain matrices L and K based on the nonsymmetric algebraic Riccati equation (33) returns $\binom{4}{2} = 6$ different answers, two of which are real valued. Of the real solutions,

$$L = \begin{bmatrix} 1.0085 & -0.9780 \\ -0.0139 & 0.2664 \end{bmatrix}, \quad K = \begin{bmatrix} 0.1273 & -2.0544 \\ -0.4303 & 1.4190 \end{bmatrix}, \tag{68}$$

provide the smallest reachable set according to Lemma 1.

Using (63) we calculate the security metric based on the set of states reachable by a stealthy zero-alarm attack, $\sqrt{\mathbf{tr}(Q^*)} = 20.85$. Of the observer and controller gain matrices that achieve γ^*, here we have selected the pair with the smallest value of $\sqrt{\mathbf{tr}(Q^*)}$. This point is plotted in Figure 1 with a blue asterisk ($*$). From Lemma 3, the open loop ICB $\|H\|_2$ gain is $\gamma_0 = 103.63$ (plotted with a red asterisk).

In Figure 1, the solid, dotted, and dashed curves present the trade-off between performance and security as determined by Theorem 3, Theorem 2, and Theorem 2 of [26], respectively. These curves are computed by solving these optimizations repeatedly for different values of $\bar{\gamma} \in [\gamma^*, \gamma_0]$ over the trade-off interval (hence each point corresponds to different solution of gain matrices). We use (63) as the security metric in all cases. The accuracy of the nonlinear optimization produced by the geometric approach reveals the true, fundamental trade-off between performance and security, highlighting the fact that giving up a small amount of $\|H\|_2$ performance can lead to a dramatic improvement in our reachability-based security metric. This is evidenced by the relatively large vertical drop in moving from $\bar{\gamma}$ from γ^* to $\gamma^* + \epsilon$ for some small $\epsilon > 0$.

The flatness of the trade-off curve near open loop performance (near γ_0) shows that there is diminishing returns on reducing performance to gain security. The nearness of all three approaches in this flat region suggest that the constraints of the optimization are relatively loose, such that the optimization problem is less challenging to solve. In contrast, closer to γ^* the optimization landscape is quite complex. For the nonlinear

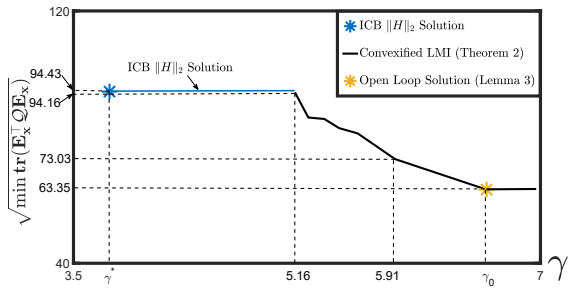

Figure 2: Performance-security trade-off for a larger system.

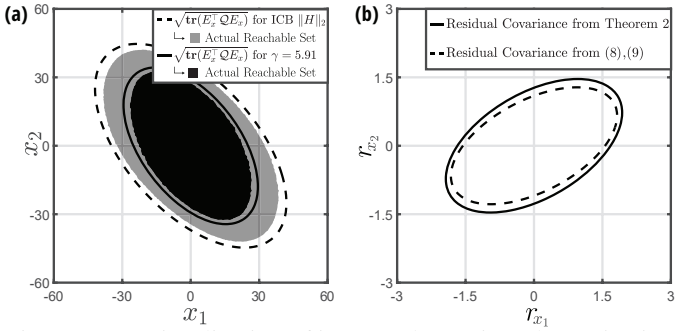

Figure 3: (a) Visualization of increased security. The projection of the optimal ellipsoids that outer-bound the reachable set, corresponding to the $\|H\|_2$ optimal gains (dashed) and secured gains (solid) with $\gamma = 5.91$. (b) The closeness of these ellipsoidal projections (solid - approximate; dashed - actual) demonstrates our linearizations preserve the structure of the original problem.

optimization of Theorem 3, we were able to find solutions down to $\bar{\gamma} = 2.53$. The linearizations needed to convexify Theorem 2 prohibits finding solutions over the entire trade off interval. In this case, the problem is infeasible on the interval $\gamma \in [\gamma^*, 32.03]$. In this interval, a feasible solution cannot be found that satisfies the manifold $\mathbf{P}_{\hat{x}} = \mathbf{P}_{x\hat{x}}$ as well as satisfying the constraints for the system under attack (LMI $\mathcal{H}_L$) and without attack (LMI $\mathcal{C}_L$).

B. Scalability

Although the geometric approach of Theorem 3 provides a highly accurate solution, the scalability of the approach is highly dependent on the system characteristics (e.g., increasing the system dimension increases the number of simultaneous nonlinear equations that need to be solved). The convex optimization of Theorem 2 offers an approach that scales well, largely due to the scalability of the optimization solver.

In Figure 2, given the gains K and L, we use $\sqrt{\mathrm{tr}(\hat{E}_x^\top \mathcal{Q} \hat{E}_x)}$ from Lemma 1 as the security metric. This figure provides the trade-off between the security and performance for a system with 12 states and 8 sensors - see (69). Applying Theorem 1, $\gamma^* = 3.77$ and this $\|H\|_2$ optimal solution corresponds to a security metric of 94.43. Based on Lemma 3, the open loop (or worst-case optimal) $\|H\|_2$ gain is $\gamma_0 = 6.60$, corresponding to a security metric of 63.35. Since no attacks can cause impact in open loop, this quantifies the contribution due to the noise only. This term is important to retain so that the attack-induced ellipsoidal contributions can leverage the geometry of the noise-induced reachable set to amplify the overall impact. This is, in part, facilitated by the fact that we incorporate the actual residual covariance in the optimization (see Remark 5). Theorem 2 is infeasible over the interval $\gamma \in [3.77,\ 5.16]$. Here we keep the result from Theorem 1 since it satisfies the performance criteria.

Figure 3a plots the first projection of the ellipsoidal bound from Lemma 1 for the $\|H\|_2$ optimal solution (Theorem 1, $\gamma = \gamma^*$) and when security has been optimized (Theorem 2) with the performance constraint $\gamma = 5.91$. This plot shows the improvement in security achieved by our methods by reducing the set of states that an attacker is able to reach through a stealthy attack. This plot also shows the actual reachable set to illustrate the tightness of the outer bounding strategy, computed using 50 terms in the geometric sum of (54).

Recall that Theorem 2 approximates the residual covariance in order to linearize the Lyapunov equation (9). Figure 3b compares the approximate residual covariance produced by Theorem 2 ($\Sigma = \Pi^{-1}$) with the true residual covariance calculated from (8)-(9) using the gains produced by Theorem 2. We show the projection of the corresponding ellipsoids onto the first two components. The closeness of these ellipsoids demonstrates that our linearizations largely preserve the structure of the original problem, here namely Lemma 4.

VI. Conclusion

In this paper, we have developed two methods to design estimator and controller gain matrices to minimize the impact of attacks on modern control systems, subject to a minimum closed-loop performance constraint. We quantify the impact of attacks by the trace of the shape matrix of the minimum trace ellipsoidal bound that contains the reachable set of states. when the system is driven by the process noise and attacks that remain stealthy to a chi-squared detector. We choose a input covariance based (ICB) $\|H\|_2$ gain as the performance criteria, which serves to avoid trivial solutions. In doing so, we are able to use these tools to characterize the trade-off between closed-loop performance and potential attack impact. Most notably, we observe that marginal reductions in closed-loop performance can enable the system to become significantly more secure through the design of these gain matrices.

We demonstrate these tools, phrased as convex and non-linear programs, respectively, and some of their advantages. While the linearizations required to recover linear matrix inequalities are shown to produce meaningful solutions, there is an opportunity to investigate alternative approaches to linearize this challenging problem.

References

[1] Y. Mo and B. Sinopoli, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Transactions on Automatic Control*, vol. 61, pp. 2618–2624, 2016.

[2] J. Milosevic, D. Umsonst, H. Sandberg, and K. H. Johansson, "Quantifying the impact of cyber-attack strategies for control systems equipped with an anomaly detector," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 331–337.

[3] N. Hashemi, C. Murguia, and J. Ruths, "A comparison of stealthy sensor attacks on control systems," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 973–979.

$$F=\begin{bmatrix}
0.801 & 0 & -0.606 & 0.078 & -0.359 & 0 & 0 & 0.025 & 0 & -0.504 & -0.394 & 0.158\\
0 & 0.109 & 0 & -0.009 & 0.159 & -0.753 & 0 & 0.012 & -0.906 & 0 & 0 & -0.226\\
0 & 0.246 & 0 & 0 & -0.092 & 0.488 & 0.209 & -0.030 & 0.502 & 0 & 0 & 0\\
0.021 & 0 & -0.001 & 0.532 & 0.035 & -0.039 & -0.042 & -0.027 & 0 & 0 & 0.066 & -0.017\\
0.079 & 0 & -0.092 & 0 & 0.392 & 0 & 0 & -0.066 & 0.289 & 0 & -0.102 & -0.064\\
0 & 0.355 & -0.155 & -0.170 & 0.059 & 1.031 & 0.072 & 0 & 0.563 & -0.611 & 0 & 0\\
0 & -0.003 & -0.005 & -0.196 & 0 & -0.121 & 0 & 0 & -0.145 & 0.018 & 0 & -0.074\\
0 & 0.191 & -0.175 & 0 & 0 & 0.752 & 0 & 0 & 0.626 & -0.452 & -0.182 & 0.131\\
0 & -0.282 & 0 & 0 & 0 & 0 & 0 & -0.107 & 0 & 0.197 & 0 & 0\\
0 & -0.080 & 0 & 0 & 0 & -0.049 & -0.310 & -0.043 & 0 & 0 & 0.065 & 0.240\\
0 & -0.302 & 0 & 0.231 & -0.178 & -0.927 & -0.385 & 0.129 & 0 & 0.625 & 0 & 0\\
-0.108 & 0 & 0 & -0.098 & 0 & 0 & 0 & 0 & 0 & 0.409 & 0 & 0
\end{bmatrix},\quad
G=\begin{bmatrix}
0.342 & 0.803 & 0.542 & 0.093 & 0.722 & 0.964 & 0.655 & 0.897 & 0.088 & 0.662 & 0.865 & 0.825\\
0.758 & 0.537 & 0.934 & 0.429 & 0.780 & 0.220 & 0.156 & 0.893 & 0.037 & 0.525 & 0.486 & 0.558\\
0.105 & 0.911 & 0.434 & 0.398 & 0.724 & 0.306 & 0.125 & 0.826 & 0.645 & 0.684 & 0.983 & 0.632\\
0.816 & 0.935 & 0.432 & 0.991 & 0.691 & 0.740 & 0.724 & 0.362 & 0.001 & 0.863 & 0.143 & 0.629\\
0.978 & 0.428 & 0.170 & 0.689 & 0.938 & 0.748 & 0.520 & 0.966 & 0.453 & 0.377 & 0.126 & 0.121\\
0.424 & 0.269 & 0.010 & 0.607 & 0.282 & 0.166 & 0.840 & 0.318 & 0.726 & 0.182 & 0.059 & 0.825\\
0.807 & 0.513 & 0.978 & 0.130 & 0.465 & 0.122 & 0.099 & 0.710 & 0.476 & 0.838 & 0.801 & 0.203\\
0.187 & 0.526 & 0.449 & 0.137 & 0.354 & 0.152 & 0.446 & 0.964 & 0.989 & 0.757 & 0.327 & 0.653\\
0.984 & 0.343 & 0.234 & 0.779 & 0.608 & 0.573 & 0.729 & 0.834 & 0.725 & 0.522 & 0.972 & 0.835\\
0.8 & 0.650 & 0.819 & 0.304 & 0.419 & 0.901 & 0.560 & 0.824 & 0.782 & 0.343 & 0.424 & 0.724\\
0.468 & 0.676 & 0.224 & 0.732 & 0.454 & 0.387 & 0.876 & 0.806 & 0.131 & 0.359 & 0.226 & 0.889\\
0.744 & 0.348 & 0.208 & 0.540 & 0.470 & 0.740 & 0.489 & 0.845 & 0.483 & 0.216 & 0.705 & 0.733
\end{bmatrix},$$

$$C=\begin{bmatrix}
0.947 & 0.051 & 0.203 & 0.571 & 0.296 & 0.915 & 0.25 & 0.410 & 0.309 & 0.452 & 0.677 & 0.860\\
0.784 & 0.988 & 0.199 & 0.010 & 0.828 & 0.319 & 0.108 & 0.397 & 0.001 & 0.325 & 0.131 & 0.714\\
0.580 & 0.958 & 0.965 & 0.449 & 0.435 & 0.507 & 0.656 & 0.014 & 0.336 & 0.011 & 0.750 & 0.910\\
0.495 & 0.265 & 0.448 & 0.701 & 0.532 & 0.006 & 0.636 & 0.208 & 0.778 & 0.386 & 0.612 & 0.550\\
0.891 & 0.580 & 0.692 & 0.670 & 0.473 & 0.705 & 0.624 & 0.311 & 0.620 & 0.936 & 0.922 & 0.491\\
0.297 & 0.158 & 0.512 & 0.949 & 0.186 & 0.171 & 0.768 & 0.385 & 0.542 & 0.801 & 0.966 & 0.540\\
0.682 & 0.385 & 0.347 & 0.785 & 0.911 & 0.821 & 0.615 & 0.305 & 0.112 & 0.930 & 0.830 & 0.704\\
0.999 & 0.950 & 0.130 & 0.135 & 0.665 & 0.135 & 0.180 & 0.506 & 0.526 & 0.780 & 0.313 & 0.917
\end{bmatrix},\quad
H=\begin{bmatrix}
0.470 & 0.084 & 0.504 & 0.274 & 0.535 & 0.826 & 0.965 & 0.396 & 0.763 & 0.100 & 0.466 & 0.643\\
0.958 & 0.410 & 0.257 & 0.812 & 0.591 & 0.365 & 0.127 & 0.917 & 0.046 & 0.448 & 0.516 & 0.833\\
0.673 & 0.264 & 0.299 & 0.691 & 0.904 & 0.790 & 0.271 & 0.345 & 0.372 & 0.115 & 0.171 & 0.947\\
0.297 & 0.061 & 0.050 & 0.454 & 0.042 & 0.108 & 0.807 & 0.919 & 0.889 & 0.460 & 0.988 & 0.084\\
0.951 & 0.608 & 0.724 & 0.435 & 0.115 & 0.417 & 0.266 & 0.764 & 0.580 & 0.226 & 0.169 & 0.257\\
0.743 & 0.190 & 0.168 & 0.780 & 0.788 & 0.358 & 0.966 & 0.791 & 0.017 & 0.071 & 0.124 & 0.909\\
0.600 & 0.107 & 0.603 & 0.682 & 0.373 & 0.593 & 0.317 & 0.514 & 0.558 & 0.241 & 0.136 & 0.837\\
0.977 & 0.333 & 0.517 & 0.506 & 0.774 & 0.063 & 0.974 & 0.925 & 0.386 & 0.176 & 0.189 & 0.213
\end{bmatrix}. \tag{69}$$

[4] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 4, pp. 370–379, 2014.

[5] C. Murguia and J. Ruths, "On model-based detectors for linear time-invariant stochastic systems under sensor attacks," *IET Control Theory & Applications*, vol. 13, no. 8, pp. 1051–1061, 2019.

[6] D. Umsonst and H. Sandberg, "Anomaly detector metrics for sensor data attacks in control systems," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 153–158.

[7] C. Murguia, I. Shames, J. Ruths, and D. Nešić, "Security metrics and synthesis of secure control systems," *Automatica*, vol. 115, p. 108757, 2020.

[8] M. A. Rotea, "The generalized h2 control problem," *Automatica*, vol. 29, no. 2, pp. 373–385, 1993.

[9] M. C. De Oliveira, J. C. Geromel, and J. Bernussou, "Extended h 2 and h norm characterizations and controller parametrizations for discrete-time systems," *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.

[10] J. C. Geromel, J. Bernussou, and M. C. De Oliveira, "H/sub 2/-norm optimization with constrained dynamic output feedback controllers: decentralized and reliable control," *IEEE Transactions on Automatic Control*, vol. 44, no. 7, pp. 1449–1454, 1999.

[11] C. G. Rieger, D. I. Gertman, and M. A. McQueen, "Resilient control systems: Next generation design research," in *2009 2nd Conference on Human System Interactions*. IEEE, 2009, pp. 632–636.

[12] L. An and G.-H. Yang, "Lq secure control for cyber-physical systems against sparse sensor and actuator attacks," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 2, pp. 833–841, 2018.

[13] L. Mili and N. V. Center, "Taxonomy of the characteristics of power system operating states," in *2nd NSF-VT Resilient and Sustainable Critical Infrastructures (RESIN) Workshop*, 2011, pp. 13–15.

[14] L. Mili and S. Shukla, "Power and communications systems as integrated cyberphysical systems," in *Proc. of 48th Allerton Conference on Communication, Control, and Computing, Allerton, IL*, 2010.

[15] M. Pajic, I. Lee, and G. J. Pappas, "Attack-resilient state estimation for noisy dynamical systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 82–92, 2016.

[16] Y. Shoukry, M. Chong, M. Wakaiki, P. Nuzzo, A. Sangiovanni-Vincentelli, S. A. Seshia, J. P. Hespanha, and P. Tabuada, "Smt-based observer design for cyber-physical systems under sensor attacks," *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 1, pp. 1–27, 2018.

[17] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[18] V. Ugrinovskii, "Distributed h_∞ estimation resilient to biasing attacks," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 458–470, 2020.

[19] M. Zhu and S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 804–808, 2013.

[20] J. Yan and C. Wen, "Resilient containment control in adversarial environment," *IEEE Transactions on Control of Network Systems*, pp. 1–1, 2020.

[21] Y. Chen, S. Kar, and J. M. Moura, "Optimal attack strategies subject to detection constraints against cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1157–1168, 2017.

[22] G. Zhu, M. Rotea, and R. Skelton, "A convergent algorithm for the output covariance constraint control problem," *SIAM Journal on Control and Optimization*, vol. 35, no. 1, pp. 341–361, 1997.

[23] F. Xu, K. H. Lee, and B. Huang, "Monitoring control performance via structured closed-loop response subject to output variance/covariance upper bound," *Journal of Process Control*, vol. 16, no. 9, pp. 971–984, 2006.

[24] K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[25] D. A. Bini, B. Iannazzo, and B. Meini, *Numerical solution of algebraic Riccati equations*. SIAM, 2011.

[26] N. Hashemi and J. Ruths, "Gain design via lmis to minimize the impact of stealthy attacks," in *2020 Annual American Control Conference (ACC)*. IEEE, 2020.

[27] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal techniques for reachability analysis of discrete-time linear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 26–38, 2007.

[28] A. A. Kurzhanskiy and Varaiya, *Ellipsoidal calculus for estimation and control*. Nelson Thornes, 1997.

[29] N. Hashemi and J. Ruths, "Generalized outer bounds on the finite geometric sum of ellipsoids," *arXiv preprint arXiv:2006.08739*, 2020.

[30] J. Lofberg, "Yalmip : a toolbox for modeling and optimization in matlab," in *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, 2004, pp. 284–289.

[31] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html

[32] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, ser. Studies in Applied Mathematics. Philadelphia, PA: SIAM, 1994, vol. 15.

[33] H.-B. Meyer, "The matrix equation az+b-zcz-zd=0," *SIAM Journal on Applied Mathematics*, vol. 30, no. 1, pp. 136–142, 1976.

Justin Ruths is an assistant professor in Mechanical Engineering and System Engineering at the University of Texas at Dallas. He received a B.S. from Rice University; M.S. from Columbia University; and Ph.D. degree in Systems Science and Applied Mathematics from Washington University in Saint Louis. His research focuses on the control and optimization of large-scale connected systems, including problems in neuroscience, cyber-physical system security, and control of networks.

Navid Hashemi is a PhD candidate at the University of Texas at Dallas. He received his B.S. from Amirkabir University of Technology, Tehran, Iran; M.S. from the University of Texas at Dallas. His research interests include cyber-physical systems, fault detection, and safe design.

APPENDIX A
PROOF OF LEMMA 1

Here $\mathcal{P}$ is the inverse of the shape matrix of the ellipsoidal bound for the ξ reachable set ($\mathcal{P}^{-1} = \mathcal{Q}$), such that the first block $E_x^\top \mathcal{Q} E_x$ is the shape matrix of the ellipsoidal bound of the reachable set of system states. To make this ellipsoidal bound as small as possible, the cost is selected to minimize the trace of the shape matrix $E_x^\top \mathcal{Q} E_x$. To use $\mathcal{Q}$ as the variable of the optimization instead of $\mathcal{P}$ we apply the transformation $T = \text{diag}\,[\mathcal{Q},\, \mathcal{Q},\, I_n]$, to (20), i.e., $T^\top \mathcal{H} T$, which results in the LMI in (17). Equation (17) is convex which implies the uniqueness of the minimum trace bounding ellipsoid.

Remark 7: Note that the parameter a is not a decision variable of the optimization in (17). It appears nonlinearly (multiplying $\mathcal{Q}$). Since a belongs to a compact interval, the conventional choice is to solve (17) across a grid search in a and select the minimal, feasible solution.

APPENDIX B
PROOF OF LEMMA 2

From (23) and the definition of ω_k, we can calculate the quadratic terms in (25),

$$h_k^\top h_k = H^\top x_k^\top x_k H + 2x_k^\top \theta_k + \theta_k^\top \theta_k, \tag{70}$$

$$\omega_k^\top \omega_k = \nu_k^\top \nu_k + \eta_k^\top \eta_k. \tag{71}$$

Taking the expectation (x_k and θ_k are independent),

$$\begin{aligned}\mathbf{E}[h_k^\top h_k] &= \mathbf{E}[H x_k x_k^\top H^\top] + \mathbf{E}[\theta_k^\top \theta_k] \\ &= \mathbf{tr}\big(H\ \mathbf{E}[x_k x_k^\top] H^\top\big) + \mathbf{tr}(R_3),\ \text{and} \qquad (72)\\ \mathbf{E}[\omega_k^\top \omega_k] &= \mathbf{E}[\nu_k^\top \nu_k] + \mathbf{E}[\eta_k^\top \eta_k] = \mathbf{tr}(R_1) + \mathbf{tr}(R_2). \qquad (73)\end{aligned}$$

The unknown quantity is then the covariance of the state, $\mathbf{E}[x_k x_k^\top]$, which is the first block of the stacked state ζ_k covariance $\mathbf{P}_k = \mathbf{E}[\zeta_k \zeta_k^\top]$. This covariance follows the update, evaluating $\mathbf{E}[\zeta_{k+1}\zeta_{k+1}^\top]$ with (24),

$$\mathbf{P}_{k+1} = \hat{A}\mathbf{P}_k\hat{A}^\top + \hat{R}, \qquad \hat{R} = \begin{bmatrix} R_1 & 0 \\ 0 & L R_2 L^\top \end{bmatrix}. \tag{74}$$

Because the matrix $\hat{A}$ is stable the covariance converges to a steady value $\lim_{k\to\infty} \mathbf{P}_k = \mathbf{P}$ which satisfies the Lyapunov equation (28).

APPENDIX C
PROOF OF THEOREM 1

The formula for the optimal gain γ^*, (34), comes naturally from the $\|H\|_2$ bound derived in (29). Since all other terms are constant, minimizing $\mathbf{tr}(H\mathbf{P}_x H^\top)$ is equivalent to minimizing the gain. This covariance is constrained by the Lyapunov equation in (28). Here, which is a standard technique for incorporating Lyapunov equations into convex optimizations, we replace this equality constraint with the very similar inequality,

$$\mathbf{P} - \hat{A}\mathbf{P}\hat{A}^\top - \hat{R} \succeq 0, \qquad \mathbf{P} \succeq 0. \tag{75}$$

We can now combine these two inequality constraints into one using the Schur complement [32],

$$\mathcal{C} = \begin{bmatrix} \mathbf{P} - \hat{R} & \hat{A}\mathbf{P} \\ \mathbf{P}\hat{A}^\top & \mathbf{P} \end{bmatrix} \succeq 0. \tag{76}$$

This relaxation is justified because the objective function $\mathbf{tr}(H\mathbf{P}_x H^\top)$ minimizes the decision variable $\mathbf{P}$ and drives the optimization to the bound of the inequality - hence driving the relaxed form (75) to the equality (28).

We use the following transformation to linearize $\mathcal{C}$,

$$\mathcal{C}_L = \begin{bmatrix} T_1 & \\ & T_1 \end{bmatrix}^\top \mathcal{C} \begin{bmatrix} T_1 & \\ & T_1 \end{bmatrix} = \begin{bmatrix} \mathbf{P}_L - R_L & \hat{A}_L \\ \hat{A}_L^\top & \mathbf{P}_L \end{bmatrix}, \tag{77}$$

with

$$T_1 = \begin{bmatrix} \mathbf{Q}_1 & I \\ \mathbf{Q}_{12}^\top & 0 \end{bmatrix}, \qquad \mathbf{P}^{-1} = \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \mathbf{Q}_2 \end{bmatrix}, \tag{78}$$

$$\mathbf{P}_L = T_1^\top \mathbf{P} T_1 = \begin{bmatrix} \mathbf{Q}_1 & I \\ I & \mathbf{P}_x \end{bmatrix}, \tag{79}$$

$$R_L = T_1^\top R T_1 = \begin{bmatrix} \mathbf{Q}_1 R_1 \mathbf{Q}_1 \ + \ \mathbf{Q}_{12} L R_2 L^\top \mathbf{Q}_{12}^\top & \mathbf{Q}_1 R_1 \\ R_1 \mathbf{Q}_1 & R_1 \end{bmatrix},$$

$$\hat{A}_L = T_1^\top \hat{A}\mathbf{P} T_1 = \begin{bmatrix} \mathbf{Q}_1 F + XC & Z \\ F & F\mathbf{P}_x + GY \end{bmatrix},$$

where X, Y, and Z are defined as

$$X = \mathbf{Q}_{12} L, \tag{80}$$

$$Y = K\mathbf{P}_{x\hat{x}}^\top, \tag{81}$$

$$\begin{aligned} Z = \mathbf{Q}_1 F \mathbf{P}_x + XC\mathbf{P}_x + \mathbf{Q}_1 GY + \mathbf{Q}_{12} F \mathbf{P}_{x\hat{x}}^\top \qquad (82)\\ + \mathbf{Q}_{12} GY - XC\mathbf{P}_{x\hat{x}}^\top. \end{aligned}$$

The term $\mathbf{P}_L - R_L$ can be linearized by applying a Schur complement to recover $\mathcal{C}_L$ in (31). This transformation changes the set of decision variables from ($\mathbf{P}_x$, $\mathbf{P}_{\hat{x}}$, $\mathbf{P}_{x\hat{x}}$, L, K) to ($\mathbf{P}_x$, $\mathbf{Q}_1$, X, Z, Y). The solution in these new decision variables is then used to calculate $\mathbf{P}_{x\hat{x}}$ and $\mathbf{Q}_{12}$ using (82) and the identity

$$\mathbf{P}_x \mathbf{Q}_1 + \mathbf{P}_{x\hat{x}} \mathbf{Q}_{12}^\top = I, \tag{83}$$

which comes from the first block of the definition $\mathbf{P}\mathbf{Q} = I$. The definition of Z in (82) and (83) combine to form the non-symmetric algebraic riccati equation (33), which is analytically computable and in general, has $\binom{2n}{n}$ different solutions, [33]. Finally, the gain matrices can be found by, $L = \mathbf{Q}_{12}^{-1} X$ and $K = Y \mathbf{P}_{x\hat{x}}^{-\top}$. Note that the solutions to the original Lyapunov equality (28) are a subset of the solutions of the relaxed inequality (75); so Theorem 1 characterizes all ICB $\|H\|_2$ optimal solutions. Consider the proposed change of variable is valid only if there is at least one real solution for (33) ■

APPENDIX D
REAL SOLUTIONS OF THEOREM 1

Consider $S_1, S_2 \in \mathbb{R}^{n\times n}$ are two distinct solutions of the NARE (33), where S_1 is computed numerically (e.g., gradient decent). Defining $\delta S = S_2 - S_1$, δS satisfies the NARE,

$$A_1(\delta S) + (\delta S)A_2 + (\delta S)\Gamma_1(\delta S) = 0, \tag{84}$$

where $A_1 = \Gamma_3 + S_1\Gamma_1$ and $A_2 = \Gamma_2 + \Gamma_1 S_1$. Assuming $\delta S \in \mathbb{R}^{n\times n}$ is invertible, then $W = (\delta S)^{-1}$ satisfies the

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3229774

Sylvester equation, $WA_1 + A_2W + \Gamma_1 = 0$. We know $W, A_1, A_2, \Gamma_1 \in \mathbb{R}^{n\times n}$, therefore, W has a unique real solution if A_1 and $-A_2$ do not share any eigenvalues, which implies S_2 is unique. Therefore, assuming δS is invertible, the NARE (33) has only two real solutions since if we take S_2 and try to find its unique pair, it is exactly S_1. The unique solution of this specific Sylvester equation is

$$\mathbf{vec}\{W\} = -\left((I_n \otimes A_2) + \left(A_1^\top \otimes I_n\right)\right)^{-1} \mathbf{vec}\{\Gamma_1\}, \quad (85)$$

where $\otimes$ is the Kronecker product and $\mathbf{vec}\{.\}$ reshapes a matrix to a vector.

Appendix E
Proof of Lemma 3

For both $L = 0$ or $GK = 0$, the state dynamics (21) are in open loop. When $L = 0$, the state estimate $\hat{x}_k$ converges to zero because the system is open loop stable and the open loop state dynamic becomes $x_{k+1} = Fx_k + \nu_k$. Similarly, when $GK = 0$, equation (21) simplifies to the same dynamics. With this state equation, the steady state covariance of the state, $\mathbf{P}_x$, is given by (36) and consequently the desired performance γ_0 should be the same in both cases.

Appendix F
Proof of Stability

We show that the LMI constraints $\mathcal{H}_L \succeq 0$ and $\mathcal{C}_L \succeq 0$ imply that the attacked system and nominal system are stable.

For $\mathcal{H}_L$, $\mathcal{H}_L \succeq 0$ implies $\mathcal{H} \succeq 0$ which implies $\mathcal{P} \succeq 0$ and $a\mathcal{P} - A^\top \mathcal{P} A \succeq 0$. Since $a \in [0, 1)$, $(1-a)\mathcal{P} + a\mathcal{P} - A^\top \mathcal{P} A = \mathcal{P} - A^\top \mathcal{P} A \succeq 0$. Thus, $A^\top$, and hence, A is stable.

Similarly, for $\mathcal{C}_L$, $\mathcal{C}_L \succeq 0$ implies $\mathcal{C} \succeq 0$ which implies $\mathbf{P} \succeq 0$ and $\mathbf{P} - \hat{A}\mathbf{P}\hat{A}^\top - \hat{R} \succeq 0$. Therefore, $\mathbf{P} - \hat{A}\mathbf{P}\hat{A}^\top \succeq 0$, which makes $\hat{A}$ stable as well.

Appendix G
Manifold of $\mathbf{P}_{\hat{x}}$=$\mathrm{P}_{x\hat{x}}$

We know that $\mathbf{PQ} = I$ which means,

$$\mathbf{P}_x\mathbf{Q}_1 + \mathbf{P}_{x\hat{x}}\mathbf{Q}_{12}^\top = I \quad (86)$$
$$\mathbf{P}_x\mathbf{Q}_{12} + \mathbf{P}_{x\hat{x}}\mathbf{Q}_2 = 0 \quad (87)$$
$$\mathbf{P}_{x\hat{x}}^\top\mathbf{Q}_1 + \mathbf{P}_{\hat{x}}\mathbf{Q}_{12}^\top = 0 \quad (88)$$
$$\mathbf{P}_{x\hat{x}}^\top\mathbf{Q}_{12} + \mathbf{P}_{\hat{x}}\mathbf{Q}_2 = I. \quad (89)$$

If we assume the points located on manifold ($\mathbf{P}_{\hat{x}} = \mathbf{P}_{x\hat{x}}$), based on (88) we can conclude, $\mathbf{Q}_1 = -\mathbf{Q}_{12}$ and based on (86) we have $\mathbf{P}_{\hat{x}} = \mathbf{P}_x - \mathbf{Q}_1^{-1}$. Furthermore, based on $e_k = x_k - \hat{x}_k$, we conclude $\mathbf{P}_e = \mathbf{P}_x + \mathbf{P}_{\hat{x}} - \mathbf{P}_{x\hat{x}} - \mathbf{P}_{x\hat{x}}^\top$, where on the mentioned manifold, $\mathbf{P}_e = \mathbf{P}_x - \mathbf{P}_{\hat{x}}$. Finally, we conclude, $\mathbf{P}_e^{-1} = \mathbf{Q}_1$.

Appendix H
Proof of Lemma 4

Consider that the inequality $\mathcal{H}_L$ already provides a lower bound on Π. We now add an additional upper constraint to sandwich and fully constrain Π. To that end, we provide a lower bound on Σ which is an upper bound for Π, through a relaxation on (8), i.e, $\Sigma - C\mathbf{P}_e C^\top - R_2 \succeq 0$, which applying a Schur complement, can be presented as,

$$\mathcal{X} = \begin{bmatrix} \Sigma - R_2 & C\mathbf{P}_e \\ \mathbf{P}_e C^\top & \mathbf{P}_e \end{bmatrix} \succeq 0. \quad (90)$$

Using the transformation $T_5 = \mathrm{diag}\begin{bmatrix}\Sigma^{-1} & \mathbf{P}_e^{-1}\end{bmatrix}$ gives

$$\mathcal{X}_L = T_5\mathcal{X}T_5^\top = \begin{bmatrix} \Sigma^{-1} - \Sigma^{-1}R_2\Sigma^{-1} & \Sigma^{-1}C \\ C^\top\Sigma^{-1} & \mathbf{P}_e^{-1} \end{bmatrix} \succeq 0, \quad (91)$$

and applying a second Schur complement returns

$$\mathcal{X}_L = \begin{bmatrix} \Sigma^{-1} & \Sigma^{-1}C & \Sigma^{-1}R_2 \\ C^\top\Sigma^{-1} & \mathbf{P}_e^{-1} & 0 \\ R_2\Sigma^{-1} & 0 & R_2 \end{bmatrix} \succeq 0. \quad (92)$$

Thus, replacing $\Sigma^{-1} = \Pi$ and $\mathbf{P}_e^{-1} = \mathbf{Q}_1 = \sigma_1\mathbf{Q}_{1\sigma}$ results in (49). We now have an upper bound on Π in terms of $\mathbf{Q}_{1\sigma}$, however, $\mathbf{Q}_{1\sigma}$ is not itself constrained. To do this, we relax the Lyapunov equation (9) equality to inequality,

$$\mathbf{P}_e - (F - LC)\mathbf{P}_e(F - LC)^\top - R_1 - LR_2L^\top \succeq 0. \quad (93)$$

We apply the Schur complement to receive

$$\mathcal{S} = \begin{bmatrix} \mathbf{P}_e - LR_2L^\top - R_1 & (F - LC)\mathbf{P}_e \\ \mathbf{P}_e(F - LC)^\top & \mathbf{P}_e \end{bmatrix} \succeq 0. \quad (94)$$

Using the transformation $T_6 = \mathrm{diag}\begin{bmatrix}\mathbf{P}_e^{-1} & \mathbf{P}_e^{-1}\end{bmatrix}$ gives

$$\mathcal{S}_L = T_6\mathcal{S}T_6^\top = \begin{bmatrix} \mathbf{P}_{eL} - R_{eL} & A_{eL} \\ A_{eL}^\top & \mathbf{P}_{eL} \end{bmatrix} \succeq 0, \quad (95)$$

where,

$$\mathbf{P}_{eL} = \mathbf{P}_e^{-1} = \mathbf{Q}_1, \quad (96)$$
$$R_{eL} = \mathbf{P}_e^{-1}(R_1 + LR_2L^\top)\mathbf{P}_e^{-1} = \mathbf{Q}_1R_1\mathbf{Q}_1 + XR_2X^\top,$$
$$A_{eL} = \mathbf{P}_e^{-1}(F - LC) = \mathbf{Q}_1F + XC.$$

Applying the Schur Complement again yields

$$\mathcal{S}_L = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_1F + XC & \mathbf{Q}_1R_1 & XR_2 \\ (\mathbf{Q}_1F + XC)^\top & \mathbf{Q}_1 & 0 & 0 \\ R_1\mathbf{Q}_1 & 0 & R_1 & 0 \\ R_2X^\top & 0 & 0 & R_2 \end{bmatrix} \succeq 0. \quad (97)$$

Finally, using the transformation matrix $T_7 = \mathbf{diag}\left(\begin{bmatrix}\frac{1}{\sqrt{\sigma_1}} & \frac{1}{\sqrt{\sigma_1}} & \sqrt{\sigma_1} & \sqrt{\sigma_1}\end{bmatrix}\right)$ and applying the change of variables in (45) results in (50).

Appendix I
Proof of Theorem 2

Fixing the pair (τ, σ_1) makes (51) a convex optimization. When the optimization is solved, fixing the pair $(\mathbf{Q}_{1\sigma}, \mathbf{X}_\sigma)$ and introducing the pair (τ, σ_1) as decision variables makes the optimization convex again and guarantees not increasing the objective function. Therefore, this algorithm guarantees convergence. Algorithm 1 terminates when σ_1 converges (with threshold ϵ). We know $\mathbf{P}_{x\hat{x}} = \mathbf{P}_{\hat{x}}$ which implies $\mathbf{P}_e = \mathbf{P}_x - \mathbf{P}_{\hat{x}}$, thus according to (47), (80) and (81), we conclude: $L = -\mathbf{Q}_1^{-1}X$ and $K = Y\left(\mathbf{P}_x - \mathbf{Q}_1^{-1}\right)^{-1}$.