2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

A saturated dynamic input allocation policy for preventing undetectable attacks in cyber-physical systems

A. Cristofaro*, S. Galeani**, M.L. Corradini*

Abstract— **The design of a dynamic allocator is proposed to tackle the presence of undetectable attacks in cyber-physical systems. Generating invisible signals that are capable to maintain the control inputs proximal to the saturation limits, the attacks are naturally discovered due to constraint violation. Two different design schemes are proposed: the first one is a direct allocation method based on a simple linear programming algorithm, while the second one is a dynamic optimization problem with a barrier function. The case-study of a consensus network illustrates the applicability and the efficacy of the proposed policy.**

I. INTRODUCTION

Cyber-Physical Systems (CPS) are ubiquitous in modern society, which is characterized by networked infrastructures in many different domains. Cyber-physical systems are referred to as those systems that integrate physical processes, computational resources and communication capabilities [1]. Transportation networks, power generation and distribution networks, industrial automation systems are examples of such systems, where control of physical plant is mediated by and integrated with a wireless communication network. If this integration can improve efficiency, it nonetheless makes the system more vulnerable to attacks launched in the cyber-domain. Recent real-world attacks targeting physical plants (such as the StuxNet attack [2] in 2010 that targeted Siemens' supervisory control and data acquisition systems, and the Maroochy water bleach [3]) raised the problem of cyber-physical security, suggesting that information security mechanisms have to be complemented with specifically designed control systems, possibly resilient against attacks and/or equipped with attack monitors [4], [5]. As a matter of fact, resiliency of a control system with respect to faults and failures has been widely addressed in the past, and a vast bibliography is available on fault detection, isolation and tolerance [6] [7]. Though some methods could be partially borrowed, peculiarities of the problem of cyber-physical security need to be considered [8]. Cyber attacks can be classified as replay attacks, false data injection, stealth attacks, deception attacks depending on whether their target is the input, the state, the output or a combination of them (see [9], [5], [10], [11], [12], [13], [14] and references therein). As pointed out in [5], combined attacks might be designed to result undetectable, i.e. invisible with respect to the system measurements. The synthesis of such attacks is based on the implicit weak redundancy of the plant: since the number of inputs is larger than the number of measured outputs, a malicious signal can be designed to be aligned with the null space of the transfer function of the plant with the purpose of corrupting the system behavior while remaining undetected. Based on some control allocation results established in [15], [16], the ideas proposed in this paper arise from a *"fight fire with fire"* principle: if one is able to design and implement an input signal that does not alter the desired system performances, i.e. is invisible for the regulated output of the plant, and simultaneously keep the total control effort close to its limits, the occurrence of an undetectable attack can be revealed by the violation of input constraints with the consequent loss of invisibility. In particular, the allocator design schemes proposed in [16] for optimizing the input steady-states are revisited here to achieve a different goal: generating an invisible signal that keeps at least one component of the overall control input saturated.

The paper is structured as follows. The CPS model is given in Section II, and the general framework of weak input redundancy is analyzed in Section III. The main contributions of the paper are provided in Section IV, where the development of two different architectures for the saturated dynamic allocation policy is addressed. Finally, the case-study of a consensus network subject to attacks is reported in Section V to support and validate the theoretical results.

II. CPS MODEL

As described in [5], a large class of cyber-physical systems is well suited for being modeled by linear differential-algebraic systems. A smaller, but still considerable portion of systems admits a nonsingular LTI representation. In this regard, let us consider the following LTI model for CPS

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \\ z &= Hx + Eu \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state of the system, $u \in \mathbb{R}^m$ is the control input, $y \in \mathbb{R}^p$ is the measured output and $z \in \mathbb{R}^r$ is the regulated output. It is worth noticing that, in the above model, the state x might include both physical and cyber variables.

Inputs to the system are supposed to be subject to constraints, such as physical limits for the mechanical inputs or

* A. Cristofaro and M.L. Corradini are with the School of Science and Technology – Mathematics Division, University of Camerino, 62032 Camerino MC, Italy. email: {andrea.cristofaro,letizia.corradini}@unicam.it

** S. Galeani is with the Dipartimento di Ingegneria Civile e Ingegneria Informatica, University of Rome Tor Vergata, 00133, Italy, sergio.galeani@uniroma2.it.

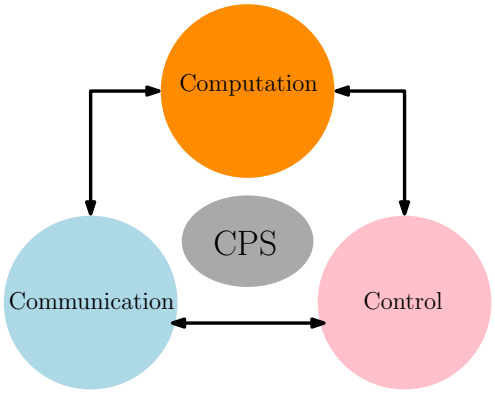

Fig. 1. Cyber-physical system sketch.

limited bandwidth for digital inputs. An overall model for the constraints is provided by the saturation operator, that is

$$u = sat_K(v), \quad sat_\kappa(\xi) = \text{sign}(\xi)\min\{|\xi|, \kappa\}$$

where $K = [k_1 \dots k_m]$ is a vector of positive numbers and the saturation has to be intended component-wise.
In the following, we will refer to any unknown, undesired and/or corrupted input signal u as an *attack* to the cyber-physical system.

III. Input redundancy

Let us set $rank[H^T \ C^T]^T = \ell \leq p + r$ and assume $\ell < m$, i.e. the dimension of the input space is larger than the dimension of the full output. This is a reasonable assumption in a real world scenario, as large scale systems are typically designed including redundancies for being prompt to handle failures and other critical conditions.
Three different plant transfer matrices are naturally defined and will be used in the following

$$\boldsymbol{W}_m(s) := (C(sI - A)^{-1}B + D) \quad (2)$$
$$\boldsymbol{W}_r(s) := (H(sI - A)^{-1}B + E) \quad (3)$$
$$\boldsymbol{W}_f(s) := ([H^T \ C^T]^T(sI - A)^{-1}B + [E^T \ D^T]^T) \quad (4)$$

where the subscripts m, r, f stand for, respectively, *measured, regulated* and *full*.

A. Output invisible signals and undetectable attacks

Definition 1: A bounded signal $w(\cdot)$ is called *partially output invisible* if it belongs to the null space of the plant transfer matrix $\boldsymbol{W}_r(s)$ from the input to the regulated output, namely if

$$(H(sI - A)^{-1}B + E)w(s) = 0 \quad \forall s \in \mathbb{C}$$

Definition 2: A bounded signal $w(\cdot)$ is called *fully output invisible* if it belongs to the null space of the plant transfer matrix $\boldsymbol{W}_f(s)$ from the input to the full output, namely if

$$([H^T \ C^T]^T(sI - A)^{-1}B + [E^T \ D^T]^T)w(s) = 0 \quad \forall s \in \mathbb{C}$$

Definition 3: An unknown bounded signal $w(\cdot)$ is called a *undetectable attack* if it belongs to the null space of the plant transfer matrix $\boldsymbol{W}_m(s)$ from the input to the measured output, namely if

$$(C(sI - A)^{-1}B + D)w(s) = 0 \quad \forall s \in \mathbb{C}$$

Undetectable attacks are finely designed for being invisible to plant sensors and simultaneously corrupt the behavior of the system with malicious scopes [5, Page 112].

B. Allocator synthesis

The realization of output invisible signals can be performed in a natural way using algebraic/geometric tools [17], [15], [18], [16], [19]. In particular, in [16] two methods have been proposed for the synthesis of annihilators, i.e. dynamic generators of output invisible signals: the first one is based on the coprime factorization of the transfer matrix and transformation to Smith form [20], [21], while the second one relies on invariant subspaces and geometric control theory [22], [23]. We will briefly sketch the rationale of the two schemes. In this regard it worth stressing that, as exploited in [16], both approaches guarantee asymptotic stability of the closed-loop plant

Algorithm 1: **Annihilator design I.**

- Let $\boldsymbol{W}(s)$ be the transfer matrix of the plant.
- Compute a left coprime polynomial factorization $\boldsymbol{D}^{-1}(s)\boldsymbol{N}(s)$ of the transfer matrix $\boldsymbol{W}(s)$.
- Compute unimodular polynomial matrices $\boldsymbol{L}(s)$ and $\boldsymbol{R}(s)$ such that $\boldsymbol{L}(s)\boldsymbol{N}(s)\boldsymbol{R}(s)$ is in Smith form.
- Define $\boldsymbol{R}_2(s) = \boldsymbol{R}(s)\begin{bmatrix} 0 \\ I_{m-p} \end{bmatrix}$.
- Choose $\boldsymbol{\Psi}(s) = \text{diag}(\psi_1(s), \dots, \psi_{m-p}(s))$ such that each $\psi_h(s)$ has real coefficient, degree not smaller than the highest degree in the $h-$th column of $\boldsymbol{R}_2(s)$, and all roots in a desired set $\mathbb{C}_g \subset \{s : \text{Re}(s) < 0\}$.
- Define the annihilator as any minimal realization of $\boldsymbol{W}_{An}(s) = \boldsymbol{R}_2(s)\boldsymbol{\Psi}^{-1}(s)$. ■

Algorithm 2: **Annihilator design II.**

- Let $\mathcal{R}^*$ be the *largest controlled-invariant subspace* of the plant.
- Define invertible matrices $T \in \mathbb{R}^{n\times n}$ and $N \in \mathbb{R}^{m\times m}$ where $T^{-1} = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ with $\text{Im}\,(T_1) = \mathcal{R}^*$ (and T_2 is such that $\det(\begin{bmatrix} T_1 & T_2 \end{bmatrix}) \neq 0$), and $N^{-1} = \begin{bmatrix} N_1 & N_2 \end{bmatrix}$ with $\text{Im}\,(N_1) = B^{-1}\mathcal{R}^*$ and $\det(\begin{bmatrix} N_1 & N_2 \end{bmatrix}) \neq 0$.
- Define new coordinates $\bar{x} = Tx$, $\bar{u} = \begin{bmatrix} \bar{u}_1' & \bar{u}_2' \end{bmatrix}' = Nu$ yielding the following block structure:

$$\bar{A} = TAT^{-1} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix},$$
$$\bar{B} = TB = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{12} \\ 0 & \bar{B}_{22} \end{bmatrix},$$
$$\bar{C} = CT^{-1} = \begin{bmatrix} 0 & \bar{C}_2 \end{bmatrix}$$

- Compute a friend $\bar{F} = \begin{bmatrix} \bar{F}_{11} & 0 \\ \bar{F}_{21} & \bar{F}_{22} \end{bmatrix}$ of $\mathcal{R}^*$ such that the spectra of $\bar{A}_{11}^F := (\bar{A}_{11} + \bar{B}_{12}\bar{F}_{21} + \bar{B}_{11}\bar{F}_{11})$ and $\bar{A}_{22}^F := (\bar{A}_{22} + \bar{B}_{22}\bar{F}_{22})$ both lie in $\mathbb{C}_g$.

- Define the annihilator as:

$$\dot{x}_{An} = \bar{A}_{11}^F x_{An} + \bar{B}_{11}\hat{u}, \tag{6a}$$

$$u_{An} = N^{-1}\left(\begin{bmatrix}\bar{F}_{11}\\ \bar{F}_{21}\end{bmatrix} x_{An} + \begin{bmatrix} I_{m-p}\\ 0\end{bmatrix}\hat{u}\right), \tag{6b}$$

with $x_{An}(0) = 0$. ■

For further reasoning, let us introduce the following function sets:

$$\mathcal{V}_\sharp := \{v : \boldsymbol{W}_\sharp(s)v(s) = 0\}, \quad \sharp \in \{m, r, f\},$$

corresponding, respectively, to the space of undetectable attacks, partially output invisible signals and fully output invisible signals.

IV. Saturated dynamic input allocation

Suppose that a nominal input u_c is commanded to the system, with the aim of achieving the tracking $e_r \to 0$ with $e_r = z - r$ for some reference signal r. For the sake of simplicity, let us assume in addition that such nominal control input does not saturate, i.e.

$$u_c(t) = sat_K(u_c(t)),\ \forall t \geq 0.$$

The proposed policy arises from the naive observation that, for the synthesis of undetectable attacks to be attainable, the components of the control input, and in particular those in the null space of the transfer matrix (2), must be not fully saturated. By converse reasoning, if one is able to artificially saturate the input signals, by possibly introducing virtual saturation limits with safety margins, this might prevent the occurrence of undetectable attacks. In other words, we are interested in the following allocation problem.

Saturated dynamic allocation problem (SDAP): Select an arbitrary small positive parameter $\epsilon > 0$ and set $K_\epsilon := K - \epsilon\mathbb{1}$, where the symbol $\mathbb{1}$ stands for the one-vector in $\mathbb{R}^m$. Find a fully output invisible signal $\upsilon(t)$ such that

$$\upsilon(t) = \text{argmin}_{\{v \in \mathcal{V}_f : \|u_c(t)+v\|_1 \leq K_\epsilon\}} \delta(K_\epsilon, u_c(t) - v)$$

where $\delta(\cdot,\cdot)$ is a distance measure to be specified, possibly having a variable structure and including logic. It might be useful to consider also the relaxed counterpart of the optimization problem, in which invisibility is required only with respect to the regulated output.

Relaxed saturated dynamic allocation problem (RSDAP): Find a fully output invisible signal $\upsilon(t)$ such that

$$\upsilon(t) = \text{argmin}_{\{v \in \mathcal{V}_r : \|u_c(t)+v\|_1 \leq K_\epsilon\}} \delta(K_\epsilon, u_c(t) - v).$$

Having solved the optimization scheme SDAP (or RSDAP), the following statement can be given.

Theorem 1 (Unveiling of undetectable attacks): Assume that an annihilator u_{An} has been be designed as the solution to SDAP (or RSDAP), and that $u_{j^\star} = \pm(K_{\epsilon,j^\star} - \eta)$ for some $j^\star \in \{1, ..., m\}$ and $\eta \geq 0$, with $u = u_c + u_{An}$ and output response $y_{An}(t) \equiv y_o(t)$, being $y_o(t)$ the nominal response of the system. Let u_{Atk} be an undetectable attack entering the system, i.e. $\boldsymbol{W}_m(s)u_{Atk}(s) \equiv 0$ and $y_{Atk}(t) \equiv y_o(t)$. Then, as long as the excitation conditions $|u_{Atk,j^\star}| \geq \epsilon + \eta$ and $\text{sign}(u_{Atk,j^\star}) = \text{sign}(u_{j^\star})$ are satisfied, the saturation limits are violated by the $j^\star$-th component of the total input $u = u_c + u_{An} + u_{Atk}$ and the undetectable attack is revealed, i.e. the overall measured output $y_{An+Atk}(t)$ verifies $y_{An+Atk}(t) \neq y_o(t)$. □

We will now present two methods to attain the minimum in the optimization problem SDAP (or RSDAP). The first one is founded on static optimization and can be applied for output regulation at constant set-points and under steady-state plant conditions, while the second one is a gradient descent method with a barrier function.

A. Direct allocation approach

Assume that the (steady-state) nominal control is a pure feedforward $\bar{u}_c = Gr$ for some fixed constant reference r, where G is a given gain matrix. Accordingly, one can feed a constant input to the annihilator in order to generate an output invisible signal with a constant steady-state. Referring to Algorithm 1, such steady-state inputs are expressed by $\boldsymbol{W}_{An}(0)\varpi$ where ϖ is an arbitrary constant input.

Assumption 1: The matrix $\boldsymbol{W}_{An}(0)$ has a uniform[1] sub-rank $\kappa \leq n_A$, that is the submatrix formed by κ arbitrary rows from $\boldsymbol{W}_{An}(0)$ has rank κ [24, Definition 2].

Let us consider a permutation $\sigma : \{1, 2, ..., m\} \to \{1, 2, ..., m\}$ and assign priorities to the control components according to

$$\sigma(1),\ \sigma(2),\ \cdots, \sigma(m).$$

For fixed σ let us denote by $\delta_\sigma \in \mathbb{R}^\kappa$ the vectors

$$\delta_\sigma = \begin{bmatrix} (-1)^{\flat_1} K_{\epsilon,\sigma(1)} + \bar{u}_{c,\sigma(1)} \\ (-1)^{\flat_2} K_{\epsilon,\sigma(2)} + \bar{u}_{c,\sigma(2)} \\ \vdots \\ (-1)^{\flat_\kappa} K_{\epsilon,\sigma(\kappa)} + \bar{u}_{c,\sigma(\kappa)} \end{bmatrix} \tag{7}$$

where $\flat = [\flat_1 \ \cdots \ \flat_\kappa]$ is a random vector in $\{0,1\}^\kappa$. Accordingly, denoting with $\boldsymbol{W}_j,\ j = 1..., m$, the rows of $\boldsymbol{W}_{An}(0)$, let us define

$$\boldsymbol{W}_{An,\sigma} = \begin{bmatrix} \boldsymbol{W}_{\sigma(1)} \\ \boldsymbol{W}_{\sigma(2)} \\ \vdots \\ \boldsymbol{W}_{\sigma(\kappa)} \end{bmatrix}$$

A steady-state input whose entries $\sigma(1), \sigma(2), ..., \sigma(\kappa)$ are fully saturated is then obtained setting $\varpi_\sigma = \boldsymbol{W}_{An,\sigma}^\dagger \delta_\sigma$.
It is worth to note that, however, such an input is not necessarily admissible since the remaining entries might overpass the saturation limits.

[1]This assumption is not strictly necessary, it is only given to ease the statement of results and simplify the notation.

Let us consider the allocator realization

$$\dot{x}_{An} = Fx_{An} + R\varpi_\sigma$$
$$u_{An} = \alpha(Zx_{An} + S\varpi_\sigma)$$

with $W_{An}(s) = Z[sI - F]R + S$ and where α is a time-varying scaling factor to be computed. In this regard, the overall control input is

$$u = u_c - \alpha u_{An} \tag{8}$$

and hence, in order to guarantee saturation achievement and feasibility (at least in steady-state conditions), one can define the optimal $\alpha = \alpha^*$ according to the scheme:

$$\alpha^* = \min_{j=1,2,\dots,m} \frac{|\bar{u}_{c,j} + \text{sign}(\bar{u}_{An,j})K_{\epsilon,j}|}{|\bar{u}_{An,j}|}, \tag{9}$$

where $\bar{u}_{An}$ denotes the steady-state of the allocator.

B. Gradient-like optimization with a barrier function

Barrier function is a well established tool for dealing with inequality constraints. On the other hand, the problem considered in this paper is not a standard one: indeed, on the one hand we are interested in keeping the control input within the saturation limits, but on the other hand our objective is to guarantee that at least one of the control components is sufficiently close to the boundary of the admissible set. A possible solution for such goal is shaping the barrier functions as double-well potentials. To this end, fix $0 < \zeta < \epsilon$ such that the chain of inequalities $K > K_\zeta > K_\epsilon$ holds component-wise, with $K_\zeta = K - \zeta\mathbb{1}$. For any $j \in \{1, ..., m\}$, let us consider a function $\mathcal{J}_j : (-K_{\zeta,j}, K_{\zeta,j}) \to [0, +\infty)$ satisfying the following properties:

- $\mathcal{J}_j$ is continuously differentiable in $(-K_{\zeta,j}, K_{\zeta,j})$
- $\mathcal{J}_j(z) = \mathcal{J}_j(-z)$ for any $z \in (-K_{\zeta,j}, K_{\zeta,j})$
- $\mathcal{J}_j(K_{\epsilon,j}) = \mathcal{J}_j(-K_{\epsilon,j}) = 0$
- $\mathcal{J}_j'(K_{\epsilon,j}) = \mathcal{J}_j'(-K_{\epsilon,j}) = 0$
- $\mathcal{J}_j(z) > 0$ for any $z \neq -K_{\epsilon,j}, K_{\epsilon,j}$
- $\mathcal{J}_j(0) = \beta_j > 0,\ \mathcal{J}_j'(0) = 0,\ \mathcal{J}_j''(0) < 0$
- $\lim_{z \to \pm K_{\zeta,j}^{\mp}} \mathcal{J}_j(z) = +\infty$

An example of function fulfilling the above properties is illustrated in Figure 2. The overall barrier function can be defined as the juxtaposition of the individual $\mathcal{J}_j$, this yielding

$$\mathcal{J}(z) = \sum_{j=1}^{m} \mathcal{J}_j(z) \tag{10}$$

with $z \in (-K_{\zeta,1}, K_{\zeta,1}) \times \cdots \times (-K_{\zeta,m}, K_{\zeta,m})$. Accordingly, referring to the annihilator realization (6), the allocator input can be optimized imposing the updating law

$$\dot{\hat{u}} = -\gamma \nabla \mathcal{J}(u_c + W_{An}(0)\hat{u}),\ \gamma > 0. \tag{11}$$

As proved in [16], the rate convergence $\gamma > 0$ can be tuned arbitrarily without affecting the output response of the plant.[2] Assigning different priorities to the control coefficients can be achieved by suitably selecting the coefficients β_j: in fact, such value is an indicator of the penalizing cost associated to the j^{th}-component of the input being far from saturation levels $\pm K_{\epsilon,j}$. A careful tuning of these parameters might be beneficial, as the cost function is not convex and thus convergence to minima is not guaranteed in principle. On the other hand it is worth to recall that, for the scheme to be successful, it is enough to obtain at least one saturated entry in the control input. Indeed, by standard Lyapunov methods, as long as a minimum exists, its local attractivity can be proved. As a matter of fact, the stationary point in 0 is intrinsically unstable, and hence convergence to minima can be expected for almost every initial condition (see for instance [25]). A detailed analysis of convergence analysis goes beyond the scopes of the paper, and therefore will not be addressed here in deep.

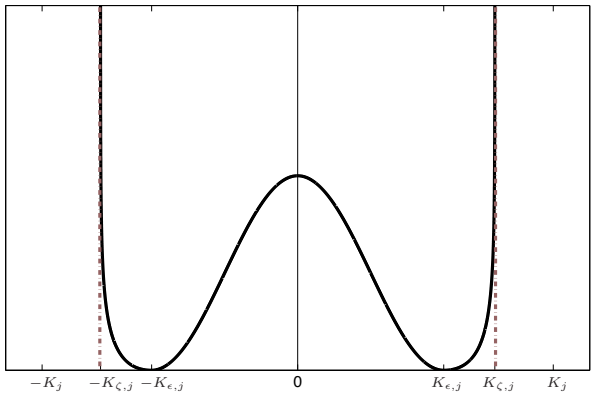

Fig. 2. Barrier function with double-well.

C. Remarks and discussion

Referring to the design scheme of Section IV-A, a strategical advantage might be changing periodically or randomly the priority permutation σ. In fact, introducing a switching scheme in the priority selection can prevent counter-actions performed by the attacker, i.e. priority reconfiguration allows to reveal the presence of clever attacks that attempt to drive the inputs away from the saturation limits.

Another fundamental issue to be considered in the allocator synthesis is the behavior of actuators transient. In this regard, we observe that both the direct approach of Section IV-A and the barrier function based approach of Section IV-B are demanded to manipulate the steady-state response of the allocator, while the transient is unconstrained to some extent. As a consequence, the response of the overall control inputs might be characterized by bumps and peaks that may potentially lead to false alarms if the saturation limits are crossed. Several techniques can be used to avoid such conditions. For instance, the allocator design might be coupled with a reference governor to guarantee that constraints are never violated while good tracking performances of the desired response are kept [26], [27], [28], [29]. Alternatively, using the results of [30], [31], [32] based on spectral decompositions, the synthesis of the allocator can be modified in order to

[2] As long as the input u_c depends only on the system output vector for which invisibility is enforced.

avoid overshoots and undershoots. However, as shown in the simulation test proposed in the next section, a clever tuning of the allocator based on practical principles, e.g. shaping an "overdamped" response, might be sufficient to bypass overshoots with a modest effort without altering the allocator structure or introducing additional mechanisms.

V. Example: a consensus network

Let us consider a 4-nodes consensus network described by an undirected graph. The natural behavior of the network is modeled by the system

$$\dot{x} = Ax, \tag{12}$$

where $x \in \mathbb{R}^4$ is the vector of agents' states and A is the Laplacian matrix of the graph

$$A = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -3 & 1 \\ 1 & 0 & 1 & -2 \end{bmatrix}$$

A measured and a regulated output have been considered, given respectively by

$$y = x_1 - x_2 - x_3 + x_4, \quad z = x_4.$$

The nodes x_1, x_2, x_3 are allowed to receive inputs u_j with a limited bandwidth $|u_j| \leq 1$ for $j = 1, 2, 3$, this corresponding to the controlled plant $\dot{x} = Ax + Bu$ with

$$B = \begin{bmatrix} I_{3\times 3} \\ 0_{1\times 3} \end{bmatrix}.$$

The nominal control input u_c is a piecewise constant signal, with reference switchings at $t = 35s$ and $t = 70s$. We notice that the open-loop plant (12) is not asymptotically stable but only marginally stable.

An allocator has been designed with the aim of saturating the inputs to reveal invisible attacks, based on the direct approach described in Section IV-A: the resulting input signal u_{An} is invisible for both the measured and the regulated output (see the black and the green plots in Figure 3 and Figure 4). The saturation limits have been set as $K_\epsilon = 0.95$ for each control component. In Figure 7 the achievement of the allocation objective is clearly visible, i.e. the control component u_3 is close to the saturation upper limit. Regarding the tuning of the allocator, the poles have been placed on the real axis with the aim of reducing the likelihood of over-and-undershoots occurrence as well as for minimizing their amplitude. In particular, the null space of the transfer matrix $\boldsymbol{W}_f(s)$ is spanned by the vector

$$\mathbf{N}^\perp(s) = [2 + s \quad 4 + 2s \quad -6 - s]^T,$$

and the annihilator has been obtained as a minimal realization of $\boldsymbol{R}_2(s)\boldsymbol{\Psi}^{-1}(s) = \mathbf{N}^\perp(s)/g(s)$ according to Algorithm 1 with

$$g(s) = (s + 1.5)(s + 5.5).$$

By this choice, the residuals in each component of the annihilator output turn out to have the same sign, and therefore the avoidance of over-and-undershoots is granted.

A sinusoidal attack has been synthesized in order to produce invisible signals for the measured output and simultaneously corrupt the response of the regulated output. As clearly illustrated in Figure 3, the output response of the system without allocation, with saturated allocation in nominal conditions and subject to the attack coincide. We notice that, in the latter case, none of the control component is saturated (see Figures 5-7). However, when the considered allocation policy is switched on, the presence of the attack is promptly revealed: the attack is no longer invisible for the measured output y due to saturation constraint violation (see the blue plots in Figure 3 and Figure 7).

VI. Conclusions

We have illustrated the design of a dynamic input allocator to deal with undetectable attacks in cyber-physical systems. The proposed architecture lies on the simple yet powerful idea that, generating invisible signals that are capable to maintain the control inputs proximal to the saturation limits, a potential attack will be naturally discovered due to constraint violation. Future investigation in this field will be devoted to enhance the design method by considering more general CPS models. e.g. differential-algebraic systems, as well as decentralized schemes for the implementation in large-scale and networked systems.

References

[1] R. Poovendran, K. Sampigethaya, S. K. S. Gupta, I. Lee, K. V. Prasad, D. Corman, and J. Paunicka, "Special issue on cyber-physical systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 1–12, 2012.

[2] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *Proc. 37th Annu. Conf. IEEE Industrial Electronics Society*, Melbourne,, 2011, pp. 4490–4494.

[3] J. Slay and M. Miller, *Lessons learned from the maroochy water breach*. Springer, 2007.

[4] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for Cyber-Physical Systems under adversarial attacks," *IEEE Trans. Autom. Contr.*, vol. 59, no. 6, pp. 1454–1467, 2014.

[5] F. Pasqualetti, F. Dorfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 110–127, 2015.

[6] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Norwell, MS: Kluwer Academic Publishers, 1999.

[7] R. Isermann, *Fault-Diagnosis Systems*. Springer, 2006.

[8] C. Aubrun, D. Sauter, and J. Yamé, "Fault diagnosis of networked control systems," *Int. J. Appl. Math. Comput. Sci.*, vol. 18, no. 4, pp. 525–537, 2008.

[9] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Contr.*, vol. 58, no. 11, pp. 2715–27 129, 2013.

[10] H. Sandberg, S. Amin, and K. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Systems*, vol. 35, no. 1, pp. 1–12, 2015.

[11] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of fault and attacks including false data injection attack in smart grid using kalman filter," *IEEE Trans. Contr. Networked Sys.*, vol. 1, no. 4, pp. 370–379, 2014.

[12] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling with energy constraint," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3023–3028, 2015.

[13] A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked control systems under cyber attacks with applications to power networks," in *American Control Conference (ACC), 2010*, 2010, pp. 3690–3696.

[14] M. L. Corradini and A. Cristofaro, "Robust detection and reconstruction of state and sensor attacks for cyber-physical systems using sliding modes," *IET Control Theory & Applications*, vol. 11, no. 11, pp. 1756–1766, 2017.

[15] A. Serrani, "Output regulation for over-actuated linear systems via inverse model allocation," in *Conf. on Decision and Control*, 2012, pp. 4871–4876.

[16] A. Cristofaro and S. Galeani, "Output invisible control allocation with steady-state input optimization for weakly redundant plants," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, 2014, pp. 4246–4253.

[17] L. Zaccarian, "Dynamic allocation for input-redundant control systems," *Automatica*, vol. 45, pp. 1431–1438, 2009.

[18] S. Galeani and G. Valmorbida, "Nonlinear regulation for linear fat plants: the constant reference/disturbance case," in *Mediterranean Control Conference*, june 2013.

[19] S. Galeani and S. Pettinari, "On dynamic input allocation for fat plants subject to multi-sinusoidal exogenous inputs," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, 2014, pp. 2396–2403.

[20] P. Antsaklis and A. N. Michel, *Linear systems*. Springer Science & Business Media, 2006.

[21] H. L. Trentelman, A. A. Stoorvogel, and M. L. J. Hautus, *Control Theory for Linear Systems*. Springer, 2001.

[22] W. Wonham, *Linear Multivariable Control. A Geometric Approach*, 3rd ed., ser. Applications of Mathematics. New York, NY: Springer Verlag, 1985, vol. 10.

[23] G. Basile and G. Marro, *Controlled and Conditioned Invariants in Linear System Theory*. Englewood Cliffs, NJ: Prentice Hall, 1992.

[24] A. Cristofaro and T. A. Johansen, "Fault tolerant control allocation using unknown input observers," *Automatica*, vol. 50, no. 7, pp. 1891–1897, 2014.

[25] J. D. Lee, M. Simchowitz, M. I. Jordan, and B. Recht, "Gradient descent only converges to minimizers," in *Conference on Learning Theory*, 2016, pp. 1246–1257.

[26] A. Bemporad and E. Mosca, "Nonlinear predictive reference governor for constrained control systems," in *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, vol. 2, 1995, pp. 1205–1210.

[27] E. G. Gilbert and I. Kolmanovsky, "Discrete-time reference governors for systems with state and control constraints and disturbance inputs," in *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, vol. 2, 1995, pp. 1189–1194.

[28] A. Bemporad, "Reference governor for constrained nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 43, no. 3, pp. 415–419, 1998.

[29] E. Gilbert and I. Kolmanovsky, "Nonlinear tracking control in the presence of state and control constraints: a generalized reference governor," *Automatica*, vol. 38, no. 12, pp. 2063–2073, 2002.

[30] R. Schmid and L. Ntogramatzidis, "A unified method for the design of nonovershooting linear multivariable state-feedback tracking controllers," *Automatica*, vol. 46, no. 2, pp. 312–321, 2010.

[31] ——, "The design of nonovershooting and nonundershooting multivariable state feedback tracking controllers," *Systems & Control Letters*, vol. 61, no. 6, pp. 714–722, 2012.

[32] L. Ntogramatzidis, J.-F. Trégouët, R. Schmid, and A. Ferrante, "Globally monotonic tracking control of multivariable systems," *IEEE Trans. on Automatic Control*, vol. 61, no. 9, pp. 2559–2564, 2016.

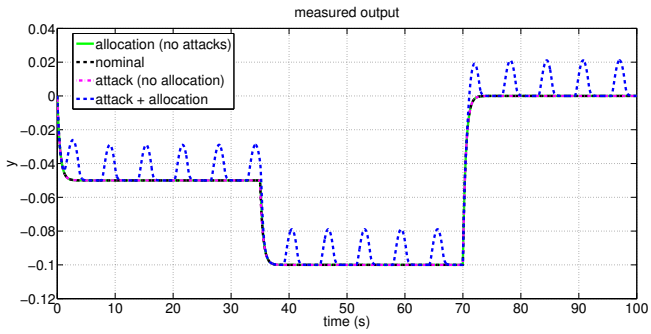

Fig. 3. Measured output $y = x_1 - x_2 - x_3 + x_4$

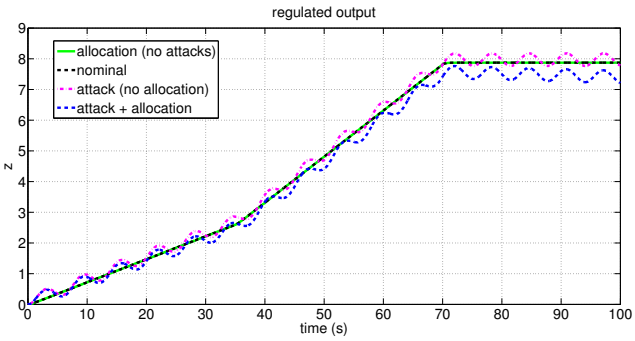

Fig. 4. Regulated output $z = x_4$

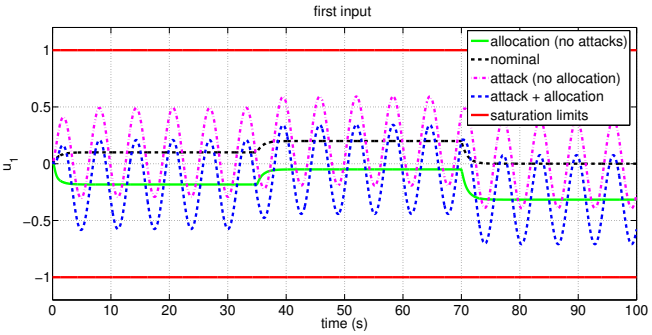

Fig. 5. Input component u_1

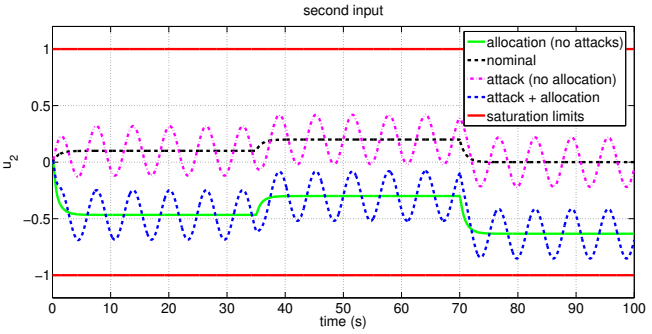

Fig. 6. Input component u_2

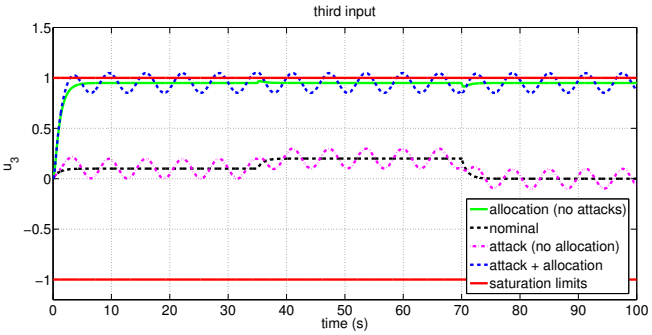

Fig. 7. Input component u_3