

Control of Average and Deviation in Large-Scale Linear Networks

Denis Nikitin, Carlos Canudas de Wit, Paolo Frasca

► **To cite this version:**

Denis Nikitin, Carlos Canudas de Wit, Paolo Frasca. Control of Average and Deviation in Large-Scale Linear Networks. IEEE Transactions on Automatic Control, Institute of Electrical and Electronics Engineers, In press, 10.1109/TAC.2021.3065191 . hal-03170606

HAL Id: hal-03170606

<https://hal.archives-ouvertes.fr/hal-03170606>

Submitted on 16 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Control of Average and Deviation in Large-Scale Linear Networks

Denis Nikitin, Carlos Canudas-de-Wit and Paolo Frasca

Abstract—This paper deals with the problem of controlling the average state of a large-scale linear network to a constant reference value. We design an output-feedback controller such that no information about state vector or system matrices is needed. For this controller to have arbitrary positive gains, it is sufficient that only a sign condition on system matrices should be satisfied. To assure that the states of the network are close to the average state, the problem of deviation minimization is solved in addition, using a novel extremum seeking algorithm.

Index Terms—Networks Control, Large-Scale Control, Extremum Seeking

I. INTRODUCTION

THE control of large-scale dynamical networks is a challenging and important question in modern control theory. Its importance originates from its multiple applications to vital systems in technology, nature and society, such as urban traffic networks [1], social networks [2], fleets of robots [3] and power networks [4]. Its difficulty originates from the large dimensionality of these real-world networks, where the number of states can reach millions. These large networks challenge the scalability of control methods from several points of view. First, the computation of traditional control algorithms becomes too expensive. Second, the structure and the detailed dynamics of the network may not be fully known. Third, the number of actuators and sensors is often much lower than the number of nodes, so that state feedback is not possible: see for instance biological neural networks, where only an average neuronal activity is measured by electrodes. Another difficulty is that the amount of energy needed to control all nodes of the network can grow unbounded as the number of nodes increases: the growth is actually exponential for some network structures [5], [6], [7]. In assessing energy-efficient control problems, it has been shown that several aspects such as network centrality [8], [9] and number of control nodes [10] play a fundamental role, but still energy analysis of large-scale network controllability is a very hard task to perform [11].

The control community has often approached the issue of network control by looking for distributed control algorithms, in which the control is applied locally at all nodes and uses only local information. Instead, in this paper we choose to work in a centralized setting, where an external operator has limited information about the network and limited access to few nodes for sensing or actuation purposes. In view of these limitations, the operator shall aim at controlling some aggregate function of the network state, rather than controlling

all of its individual nodes. A natural choice for such aggregate function is the average of the node states, which has indeed been defined as a control objective in some prior work that was motivated by opinion dynamics in social networks [12], [13]. More broadly, the control of a generic output of a large-scale network was studied in [14], [15], [16], [17], [18]: the energy required to control aggregated outputs instead of all states is much less.

In this paper, we focus on the problem of controlling the average state of the network, together with the concurrent minimization of its deviation. The average state of the network is defined as an average over all node states of the network, while the squared deviation is defined as an average over all squared differences between node states and the average state. While using a controller for the average state, it is natural to desire the system states to be close to the average: this behavior can be obtained by minimizing their squared deviation. In opposition to previous work, it is assumed that *the only values that are measured and regulated are the values of the system outputs, i.e. the average state and the squared deviation*. Moreover, we make sure that *the system model is not used in the controller*. Thus, the controller directly utilizes only system outputs and reference point and the equilibrium of internal states is never computed explicitly. Not using the system model circumvents all issues about computational complexity and uncertainties that affect large networks.

In our problem formulation we assume that the network operator has knowledge of the average and the squared deviation values. There are many physical examples of systems where the average and the squared deviation can be measured without measuring the states of the nodes. Here we briefly mention four examples:

Urban traffic networks. Consider a network of roads in a city, where the state of each node is the number of cars on the corresponding road section. The total number of cars in the city can be estimated either directly or indirectly. A direct estimation of the number of cars can be performed by vision-based methods, by processing images taken from satellites, as in [19], [20]. Although every car is counted independently, the estimation error is defined as a discrepancy in the overall number of cars, therefore these methods effectively reconstruct the total number of cars. An indirect estimation can be based on the vehicle emissions: combustion engines produce CO_2 , which then goes into the atmosphere. The polluted atmosphere changes its reflection properties based on the amount of CO_2 , thus this amount can be measured using infrared sensors mounted on satellites [21]. Therefore, the number of cars can be reconstructed from the satellite measurements. The total number of cars divided by the number of road sections equals

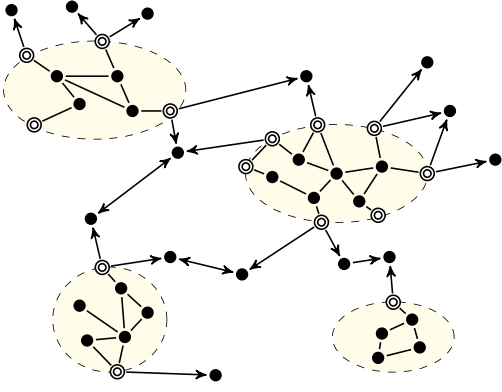


Figure 1. Scale-Free network with hubs (shaded in yellow) controlled from the boundary nodes (double circles)

to the average state of the network.

Biological neural networks. A widely known method of monitoring the brain activity is the *electroencephalography*, with electrodes placed usually along the sculp of the person being monitored. Each electrode measures voltage fluctuations of group of neurons under the surface, therefore it is directly related to the average of individual states of neurons, which obviously cannot be measured independently [22].

Dynamics of gas. Every gas consists of a huge number of particles colliding with each other, therefore it can be seen as a dynamical network with neighbouring particles whose interaction depends on their velocities. Thus, we can define the states being the velocities of each individual particle. The gas temperature can be easily measured, but at the same time it corresponds to the internal kinetic energy: $E_k = \frac{3}{2}k_B T = \frac{1}{2}mv_{rms}^2$. Here k_B is the Boltzmann constant, T is the temperature and m is the mass of one particle. The variable v_{rms}^2 represents the mean squared deviation of the velocities of particles with respect to the flow velocity. The flow velocity itself is a "wind speed", which represents the average state of the system and can be also directly measured.

Density of a fluid. Fluids also consist of a huge number of particles, and one way to write the dynamical model of a fluid is to consider a space partitioned into individual cells with states being defined as the densities of the fluid inside each cell. In this case, the average state would be the average density in the system: this density can indeed be measured for cryogenic fluids by measuring permittivity by a technique called electrical capacitance tomography [23].

Another relevant setup is the scale-free control approach to large-scale networks [24] illustrated in Fig. 1: in this approach, the goal is to control the average state and the deviation of the "hub" regions and the control is applied to the boundaries of the hubs. Output controllability of a scale-free network is studied in [25], the dual problem of reconstruction of an averaged state is solved in [26], [27], and methods to reduce a large network to a scale-free one, through the suitable aggregation of hub regions, are presented in [28], [29].

The first part of our work is devoted to the problem of controlling the average state of the linear network, where three stability results are presented. Theorem 1 provides conditions

on the integral controller gains for the stability of the closed-loop system, and Theorem 2 simplifies the conditions under the assumption that the system is positive. Then, Theorem 3 is our main contribution which gives a simple sufficient sign condition on the system matrices which guarantees stability of *any* positive integral controller for controlling the system output to a constant reference point without knowledge of the system matrices.

Most of our results regarding the output regulation problem of a large linear network system are presented under the assumption that the system is stable and positive (that is, the system matrix has positive elements outside the main diagonal). Network systems with stable dynamics and positive edge weights belong to this class. More generally, positive systems are an important class of systems for which the synthesis of large-scale control algorithms can be greatly simplified. Their impulse response is bounded by their static gain [30], optimal [31] and robust [32] feedback control laws can be easily designed using linear programming, and the state feedback output regulation problem can be explicitly solved [33]. From the passivity analysis in the classical control theory it is known that the feedback interconnection between a linear operator with an integral controller is stable irrespective of the gain (has an infinite gain margin) if the linear operator is strictly positive real (SPR) [34], [35]. From this point of view our analysis provides a new simple sufficient condition for the positive system to be SPR, which is summarized in Theorem 4.

The second part of the work focuses on the deviation minimization problem, when the system should be driven to the particular average state while the control inputs should be balanced in such a way that the squared deviation of the states takes the smallest possible value. To solve this problem we use an extremum seeking scheme [36], [37] which is an adaptive model-free algorithm for the minimization/maximization of a nonlinear steady-state output characteristic. We augment this algorithm with an additional subsystem such that both tasks are accomplished simultaneously: the average is driven to the particular value while the squared deviation is minimized. Theorem 5 proves that the system approaches any small neighbourhood of the optimum state provided the gains of the controller are small enough.

An incomplete account of this research appeared as [38]: the latter short version is much narrower than the present version as it focuses on the statement of Theorem 3 and on the interpretation of the sufficient condition therein (without providing Theorems 1, 2, 4 or 5, nor the detailed proof of Theorem 3).

Notation: Along this work several types of vector inequalities are used. If $x \in \mathbb{R}^n$, then $x \geq 0$ means $x_i \geq 0 \forall i \in \{1..n\}$, $x > 0$ means $x_i > 0 \forall i \in \{1..n\}$ and there exists $j \in \{1..n\} : x_j > 0$, and $x \gg 0$ means $x_i > 0 \forall i \in \{1..n\}$. Matrix I denotes the identity matrix of an appropriate size, while column vector of ones is denoted as $\mathbf{1}$. The set of complex numbers with non-negative real parts is \mathbb{C}^+ . We use the standard vector norm $\|x\| = \sqrt{x^T x}$ for $x \in \mathbb{R}^n$ throughout the paper.

II. PROBLEM FORMULATION

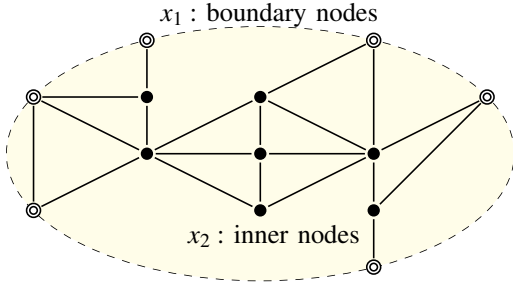


Figure 2. Network with boundary and inner nodes separation

We start the problem formulation with an example: assume the system we need to control is the network given by the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of vertices and \mathcal{E} is the set of edges. The number of vertices $|\mathcal{V}|$ is denoted by n .

On each node $v_i \in \mathcal{V}$ the state x_i is defined. Each edge $e \in \mathcal{E}$, where $e = \{v_i, v_j\}$, corresponds to the flow between nodes v_i and v_j . Matrix $A \in \mathbb{R}^{n \times n}$ represents flow ratio. The set of nodes \mathcal{V} is split into two parts \mathcal{V}_1 and \mathcal{V}_2 with state vectors x_1 and x_2 respectively (see Fig. 2). The set \mathcal{V}_1 consists of the nodes which are directly controlled by the control action u . We call these nodes "boundary". The set \mathcal{V}_2 is a set of uncontrolled nodes, which we call "inner nodes". The average state $y = \mathbf{1}^T x / n$ and the squared deviation $V = \|x\|^2 / n - y^2$ are measured. Thus the network depicted on Fig. 2 can be viewed as one particular hub from the scale-free network on Fig. 1.

The evolution of the states x , the average state y and the squared deviation V is given by the following linear time-invariant system

$$\begin{cases} \dot{x}_1 = A_{11}x_1 + A_{12}x_2 + u, \\ \dot{x}_2 = A_{21}x_1 + A_{22}x_2, \\ y = \frac{1}{n}\mathbf{1}^T x, \\ V = \|x\|^2 / n - y^2. \end{cases} \quad (1)$$

Most of real-world networks are internally stable, so we further assume A being stable. Also in Theorems 2 and 3 we assume A is a Metzler matrix (which means its off-diagonal elements are non-negative), which means all edges have positive weights. Such choice of system matrix together with the fact that $B > 0$ and $C > 0$ means that the system (1) belongs to the class of positive systems.

It is useful to analyse more general case than (1), with general stable matrix $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{1 \times n}$, $B \in \mathbb{R}^{n \times k}$, and symmetric positive semi-definite matrix $P \in \mathbb{R}^{n \times n}$ defining the quadratic output:

$$\begin{cases} \dot{x} = Ax + Bu, \\ y = Cx, \\ V = x^T Px. \end{cases} \quad (2)$$

System (1) can be written in form of (2) using

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} I \\ 0 \end{pmatrix}, \\ C = \frac{1}{n}\mathbf{1}^T, \quad P = \frac{1}{n}I - \frac{1}{n^2}\mathbf{1}\mathbf{1}^T$$

In general, the control goal is to stabilize average state y over the whole network to some desired constant state y_d without the explicit knowledge of system matrices. It is assumed that the number of states is too large that it is impossible to use full-state feedback or to use matrix A explicitly.

In the following sections two problems will be solved:

- 1) **Control of average:** find the control law $u = u(y)$ for the system (2) such that

$$\lim_{t \rightarrow \infty} y(t) = y_d.$$

- 2) **Control of average and deviation:** Find the control law $u = u(y, V)$ for the system (2) such that

$$\lim_{t \rightarrow \infty} y(t) = y_d, \quad \text{and } V \text{ is minimized.}^1$$

Section III is devoted to the first problem of average control, while Section IV shows the solution for the second problem of simultaneous control of average and deviation.

III. AVERAGE STATE REGULATION ONLY

In this section, we solve Problem 1 of controlling the average. We first describe a general controller structure and then prove stability conditions for low and high gains. Further we discuss these conditions and present examples of the controller performance.

A. Controller Structure

Define transfer function of the system (2)

$$W_s(s) = C(sI - A)^{-1}B, \quad (3)$$

thus $y(s) = W_s(s)u(s)$. Denote error between desired output and system output: $e = y_d - y$. Then we can define controller transfer function $W_c(s)$ such that $u(s) = W_c(s)e(s)$. System control loop is depicted on Fig. 3. Thus input-output

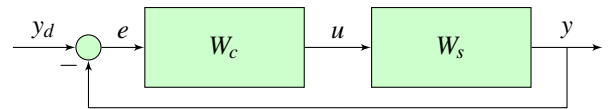


Figure 3. Control loop given by closed-loop transfer function (6)

equation is

$$y(s) = W_s(s)W_c(s)e(s), \quad (4)$$

or, solving for y ,

$$y(s) = \frac{W_s(s)W_c(s)}{1 + W_s(s)W_c(s)}y_d. \quad (5)$$

Define closed-loop transfer function

$$W(s) = \frac{W_s(s)W_c(s)}{1 + W_s(s)W_c(s)}. \quad (6)$$

In the following we investigate what properties do W_s and W_c have and what properties should W have in order to be

¹Note that the problems are formulated in the steady-state, therefore we will not pursue any optimization of the transient process.

stable. Both W_s and W_c are vector-valued, because controller input e and system output y are scalars, while u which is controller output and system input is a vector, $u \in \mathbb{R}^k$. Let us look at the i -th component of W_s and W_c , where $i \in \{1..k\}$, and define polynomials $\alpha(s), \beta_i(s), \delta(s), \gamma_i(s)$ such that

$$W_s(s)_i = \frac{\beta_i(s)}{\alpha(s)} \quad \text{and} \quad W_c(s)_i = \frac{\gamma_i(s)}{\delta(s)}. \quad (7)$$

It is obvious that $\alpha(s)$ is a polynomial of degree n . Moreover, our system is strictly stable, non-casual and matrix D (which directly connects y and u) is zero, thus

$$\deg \beta_i(s) < \deg \alpha(s) = n \quad \text{and} \quad \alpha(s) \neq 0 \quad \forall s \in \mathbb{C}^+. \quad (8)$$

We can choose $\alpha(s)$ and $\beta_i(s)$ such that $\alpha(s) \in \mathbb{R}$ for $s \in \mathbb{R}$ and $\alpha(s) > 0$ for $s \geq 0$.

Then, the controller W_c should also be stable and non-casual, which means

$$\deg \gamma_i(s) \leq \deg \delta(s) \quad \text{and} \quad \delta(s) \neq 0 \quad \forall s \in \mathbb{C}^+. \quad (9)$$

Again, it is possible to choose $\delta(s)$ and $\gamma_i(s)$ such that $\delta(s) \in \mathbb{R}$ for $s \in \mathbb{R}$ and $\delta(s) > 0$ for $s > 0$.

Now we can rewrite $W(s)$ in terms of polynomials:

$$W(s) = \frac{\sum_i \beta_i(s) \gamma_i(s)}{\alpha(s) \delta(s) + \sum_i \beta_i(s) \gamma_i(s)}. \quad (10)$$

The closed-loop transfer function $W(s)$ should have the following property: for the constant input y_d it should give the same output y , thus $W(0) = 1$. This means that $\alpha(0)\delta(0) = 0$, which is possible only if $\delta(0) = 0$, so $\delta(s)$ cannot contain free term. The simplest possible controller that satisfies this necessary condition is the integral controller given by

$$W_c(s)_i = \kappa \frac{\gamma_i}{s}, \quad (11)$$

where $\gamma \in \mathbb{R}^k$ is the vector of gains, defining relative control force applied to different actuated nodes, and κ is the overall gain. The following sections will be devoted to the integral controller and its properties.

B. Stability of the integral controller

Assume we apply the integral controller (11) to the system (2). The closed-loop system may be unstable, and in general in order to prevent this one needs to carefully choose controller gains κ and γ in (11).

Theorem 1. *The system (2) with applied integral controller (11) is asymptotically stable if $-CA^{-1}B\gamma > 0$ and $\kappa \in (0, \kappa^*)$ for some small $\kappa^* \in \mathbb{R}$.*

Proof of Theorem 1. Applying the integral controller, the transfer function of the closed-loop system is given by

$$W(s) = \frac{\kappa \sum_i \beta_i(s) \gamma_i}{\alpha(s)s + \kappa \sum_i \beta_i(s) \gamma_i}. \quad (12)$$

For the stability of the closed-loop system $W(s)$ should have no poles on the right-hand side of the complex plane \mathbb{C}^+ . Decompose the denominator:

$$\alpha(s)s \left(1 + \kappa \frac{\sum_i \beta_i(s) \gamma_i}{\alpha(s)s} \right) \neq 0.$$

Denote $Q(s) = \left(\sum_i \beta_i(s) \gamma_i \right) / (\alpha(s)s)$. Any point such that $\alpha(s) = 0$ or $s = 0$ leads to $W(s) = 1$, thus the poles of the transfer function can arise only in the roots of $1 + \kappa Q(s)$. We will prove that there exists κ^* such that

$$\forall \kappa \in (0, \kappa^*) : \forall s \in \mathbb{C}^+ \setminus \{0\} \quad \text{Re}\{1 + \kappa Q(s)\} > 0. \quad (13)$$

Choose $\varepsilon, R \in \mathbb{R}$ such that $\varepsilon < |\lambda_i(A)|$ and $R > |\lambda_i(A)|$ for all $i \in \{1..n\}$. Thus all the roots of $\alpha(s)$ lie in a ring between ε and R in a left half-plane. We split complex right half-plane \mathbb{C}^+ into three parts:

$$\begin{aligned} H_{0,\varepsilon}^+ &= \{s : \text{Re } s \geq 0, |s| < \varepsilon\} \\ H_{\varepsilon,R}^+ &= \{s : \text{Re } s \geq 0, |s| \geq \varepsilon, |s| \leq R\} \\ H_{R,\infty}^+ &= \{s : \text{Re } s \geq 0, |s| > R\} \end{aligned}$$

First we analyse $H_{0,\varepsilon}^+$. Function $Q(s)$ has a pole at zero, thus it can be written using Laurent series with coefficients Q_n :

$$Q(s) = \frac{Q_{-1}}{s} + \sum_{n=0}^{\infty} Q_n s^n = \frac{Q_{-1}}{s} + P(s),$$

where $P(s)$ is an analytic function. The residual $Q_{-1} = \left(\sum_i \beta_i(0) \gamma_i \right) / \alpha(0) = -CA^{-1}B\gamma > 0$, thus

$$\text{Re} \frac{Q_{-1}}{s} \geq 0 \quad \forall s \in H_{0,\varepsilon}^+ \setminus \{0\},$$

while $P(s)$ is analytic in \mathbb{C} and thus has a minimum in $H_{0,\varepsilon}^+$.

Next we analyse $H_{R,\infty}^+$. This set is contained into a set $H_{R,\infty} = \{s : |s| > R\}$. If R is big enough, $Q(s)$ is analytic in $H_{R,\infty}$, but it vanishes at infinity, therefore by a maximum modulus principle $Q(s)$ is bounded from below in $H_{R,\infty}$ by values on its boundary, and consequently it is bounded in $H_{R,\infty}^+$.

Finally, set $H_{\varepsilon,R}^+$ is compact and does not contain zeros or roots of $\alpha(s)$. Therefore $Q(s)$ is analytic in it and thus bounded. We obtained that $\text{Re } Q(s)$ is bounded from below in $\mathbb{C}^+ \setminus \{0\}$. Denoting this bound as Q_{inf} , we see that choosing $\kappa^* = -1/Q_{inf}$ in case $Q_{inf} < 0$ or $\kappa^* = +\infty$ in case $Q_{inf} \geq 0$ assures satisfaction of (13) and therefore proves the theorem. \square

Theorem 2. *The system (2) with applied integral controller (11) is asymptotically stable if the system (2) is positive, $\gamma > 0$, $-CA^{-1}B\gamma \neq 0$ and $\kappa \in (0, \kappa^*)$ for some small $\kappa^* \in \mathbb{R}$.*

Proof of Theorem 2. Positivity of the system (2) means that all elements of matrices B and C are greater or equal than zero, and matrix A is a Metzler matrix. Now we introduce the notion of M-matrix:

Definition 1 (M-matrix, [39]). An $n \times n$ matrix M that can be expressed in the form $M = \alpha I - L$, where $L = (l_{ij})$ with $l_{ij} \geq 0$, $1 \leq i, j \leq n$, and $\alpha \geq \lambda(L)$ where $\lambda(L)$ is the maximum of the moduli of the eigenvalues of L , is called an M-matrix.

From this definition it follows immediately that a negative of a Metzler stable matrix is an M-Matrix. The main property of any M-matrix M is that its inverse M^{-1} is a positive matrix, thus $(M^{-1})_{ij} \geq 0$ for all i, j . [40]

Matrix $-A$ is an M-Matrix which means that $-A^{-1}$ has its all elements nonnegative, therefore $-CA^{-1}B$ is a positive vector. By the theorem statement $\gamma > 0$, and having $-CA^{-1}B\gamma \neq 0$ means $-CA^{-1}B\gamma > 0$ and, applying Theorem 1, this leads to the asymptotic stability of the closed-loop system for small enough κ . \square

By Theorem 2, if the system is positive it is enough to choose $\gamma = \mathbf{1}$ (or any $\gamma \gg 0$, just to satisfy $-CA^{-1}B\gamma \neq 0$), and then pick up small enough overall gain κ .

C. Controller with arbitrary large gains

It appears although that there exists simple criteria on the system matrices which says whether the closed-loop system will converge irrespectively of controller gain values, provided that they are positive. This result is one of the main contributions of this paper, and it is formulated as follows:

Theorem 3. *The system (2) with applied integral controller (11) is asymptotically stable for arbitrary large positive controller gains κ and γ if the system (2) is positive, $CA^2 > 0$ and $CA^2B\gamma > 0$.*

This result means that, irrespectively of the gains, the integral controller will preserve stability for a very large class of systems. One of the important types of large-scale networks for which Theorem 3 is satisfied is a general consensus network (e.g. for social interactions), see the example below.

Example (Damped consensus). Assume the system (2) is given by matrices $A = -L - \alpha I$, where L is a Laplacian matrix of some network with n nodes, $\alpha > 0$ means additional damping to the system to preserve stability, and $C = \mathbf{1}^T/n$ represents average state of the network. Then A is a Metzler stable matrix, and C is the eigenvector of A with corresponding eigenvalue $-\alpha$, thus $CA^2 = \alpha^2 C > 0$. Then any controller with positive gains κ and γ will lead to the convergence, provided $B\gamma > 0$.

One should notice that the condition $CA^2B\gamma > 0$ on the control matrix B is very non-restrictive, because by choosing appropriate vector gain γ it is always possible to make $B\gamma > 0$, and hence, provided $CA^2 > 0$ and $CA^2B \neq 0$, we will have $CA^2B\gamma > 0$. The reason for this is the fact that the regulation variable is a single scalar output.

Proposition 1. *Condition $CA^2B\gamma > 0$ is a sufficient condition for the output controllability of the system (2).*

Proof. Indeed, Kalman rank test for the output controllability of (2) can be written as

$$\text{rank}\{C \begin{pmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{pmatrix}\} = 1,$$

and by $CA^2B\gamma > 0$ we have $CA^2B \neq 0$, which means that the rank test is satisfied. \square

Note that the analogue of this Proposition can be proven for Theorems 1 and 2.

Corollary 1. *The positive system (2) with $CA^2 \gg 0$ is asymptotically stable for any integral controller (11) with positive gains applied to any single boundary node. Therefore, it is enough to control only one node.*

Before proving the third theorem we need to state three technical lemmas.

Lemma 1. *Suppose we have a matrix $\mathcal{M} = M + ibI$, which is a complex matrix with real part M and imaginary part bI , with $b \in \mathbb{R}$ and I an identity matrix. Assume M being invertible and having no eigenvalues on the imaginary axis. Denote $\mathcal{L} = \mathcal{M}^{-1} = L + i\bar{L}$. Then the real part of \mathcal{L} is given by*

$$\text{Re } \mathcal{L} = L = (M + b^2M^{-1})^{-1}. \quad (14)$$

Proof. See Appendix A. \square

Lemma 2. *Let M be an M-matrix. Let C be a row-vector such that $CM^2 > 0$. Then*

$$C(M + tM^{-1})^{-1} > 0 \quad (15)$$

for any $t \geq 0$.

Proof. See Appendix B. \square

Lemma 3. *Let M be an M-matrix. Let C be a row-vector such that $CM^2 > 0$ and $CM^2B\gamma > 0$. Then*

$$C(M + tM^{-1})^{-1}B\gamma > 0 \quad (16)$$

for any $t \geq 0$.

Proof. See Appendix C. \square

Proof of Theorem 3. Applying the integral controller and multiplying nominator and denominator by s , the transfer function of the closed-loop system is given by

$$W(s) = \frac{\kappa C(sI - A)^{-1}B\gamma}{s + \kappa C(sI - A)^{-1}B\gamma}. \quad (17)$$

It is sufficient to show that the real part of denominator is strictly greater than zero in the right half-plane. Since $\text{Re } s > 0$ in the right half-plane, it is enough to show that $\text{Re}\{\kappa C(sI - A)^{-1}B\gamma\} > 0$.

Denote $\text{Re } s = \alpha$ and $\text{Im } s = \beta$, so matrix $(sI - A)^{-1} = ((\alpha I - A) + i\beta I)^{-1}$. Denote $M = \alpha I - A$. Matrix A is a Metzler stable matrix, thus $(-A)$ is an M-matrix and matrix M is an M-matrix too. Moreover, condition $CA^2 > 0$ implies $CM^2 > 0$ and $CA^2B\gamma > 0$ implies $CM^2B\gamma > 0$. Applying Lemma 1 we conclude that

$$\text{Re } \kappa C(M + i\beta I)^{-1}B\gamma = \kappa C(M + \beta^2M^{-1})^{-1}B\gamma. \quad (18)$$

By Lemma 3 $C(M + tM^{-1})^{-1}B\gamma > 0$ for any $t \geq 0$, and assuming $\kappa > 0$ we trivially obtain a sufficient condition on positivity of the real part of the denominator. \square

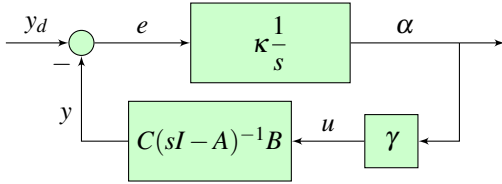


Figure 4. Feedback interconnection of passive systems

Lemmas 1-3 allow us to formulate a more general theorem, applicable to a general SISO linear system:

Theorem 4. *Any positive stable SISO system with control matrix $B \in \mathbb{R}^{n \times 1}$ and observation matrix $C \in \mathbb{R}^{1 \times n}$ such that $CA^2 > 0$ and $CA^2B > 0$ has a Strictly Positive Real (SPR) transfer function, thus it is strictly-input passive.*

Proof. The SPR transfer function should by definition satisfy $\text{Re}\{C(sI-A)^{-1}B\} > 0$ for $\text{Re}S > 0$, see [34], [35]. Therefore the proof of this theorem follows the same steps as the second part of the proof of Theorem 3. \square

From the point of view of linear systems theory, Theorem 4 is the main result of our paper. Indeed, such a simple condition for passivity for positive systems appears in literature for the first time.

Moreover, it appears that Theorem 3 is a direct consequence of Theorem 4, as the following reasoning shows. Assume we fix an input gain vector γ and define a new controller output α such that $u = \gamma\alpha$. Then the system (2) becomes SISO with respect to input variable α . Define $H_1(s) = \kappa/s$ and $H_2(s) = C(sI-A)^{-1}B\gamma$. It is possible to construct control loop with feedback interconnection as depicted on Fig. 4. The closed-loop system input is defined as y_d and the system output is α . It is known that for L_2 stability of the system with feedback interconnection it is sufficient that the transfer function of one of the blocks is *positive real* (PR, which is equivalent to passivity) and another is *strictly positive real* (SPR) [34]. Passivity of an integral controller $H_1(s)$ is obvious, and Theorem 4 is used to prove that $H_2(s)$ is SPR. Therefore the closed-loop system is L_2 stable. Now, it remains to prove that $y \rightarrow y_d$, which is obvious if one recalls that an output of a stable system with constant input converges to a constant value, thus for any constant y_d there exists α^* such that $\alpha \rightarrow \alpha^*$. But convergence of an output of an integral controller means that its input converges to zero, which reads as $e \rightarrow 0$, which is exactly $y \rightarrow y_d$.

D. Practical use of the theorems

Theorems 1-3 presented in the previous sections provide the same result, the stability of the closed-loop system (2) with controller (11). However, they differ in their assumptions. During the derivation of the controller (11) we assumed that the system matrices are not known. However, usually one has a knowledge about some general properties of the system, such as positivity. These properties can sometimes be induced from the nature of the problem itself and do not rely on the particular topology. Therefore the results presented in our work can be used to analyse the stability based on these properties.

Theorem 1 requires $-CA^{-1}B\gamma > 0$ for the integral controller to be stable for $\kappa \in (0, \kappa^*)$. This scalar condition essentially means that the direction of adaptation of integral controller forms an acute angle with the zero frequency gain of the system. In practice one usually knows the direction of the zero frequency gain. At worst, it is enough to change the sign of γ once.

Theorem 2 exploits positivity of the system: the zero frequency gain of the positive system is positive. Therefore it is enough to use positive gains for the integral controller, and the condition $-CA^{-1}B\gamma > 0$ can be loosened just to $-CA^{-1}B \neq 0$. However the gain κ still should satisfy $\kappa \in (0, \kappa^*)$.

In Theorem 3 the small-gain condition $\kappa \in (0, \kappa^*)$ is removed at the cost of adding the vector inequality $CA^2 > 0$. This inequality can be used to determine stability without knowing the particular matrices for some classes of systems, such as damped consensus, see the example after Theorem 3. For other systems, this condition should be interpreted as a constraint on the system parameters.

In the remainder of this section we will analyse the condition $CA^2 > 0$ more closely. Namely, first of all we will prove that this condition cannot be relaxed, since weaker conditions would not assure the stability for all gains. Then, we will provide some graph-theoretical intuition and rewrite the condition in terms of quadratic constraint on the node self-dampings.

If A is a Metzler stable matrix, all elements of A^{-1} are nonpositive. Multiplication of a positive vector by a matrix with nonpositive elements renders negative vector, therefore right multiplying the condition $CA^2 > 0$ by A^{-1} one obtains $CA < 0$, and the same argument provides $C > 0$. The condition $CA^2 > 0$ is new and it is used in the Lemmas 2 and 3 (substituting $M = \alpha I - A$ as in the proof of the theorem). When one looks at the statement of the Lemma 2, one might think that it would be enough to require the less restrictive condition $CA < 0$ (This condition is obtained from the statement of the Lemma 2 by letting $t \rightarrow +\infty$) and has been proposed for the full state static feedback output control of positive systems by [33]).

However, let us show that condition $CA^2 > 0$ is significant and $CA < 0$ is not sufficient. An example of the positive system with $CA < 0$ but $CA^2 \not> 0$ would be

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad C = \mathbf{1}^T. \quad (19)$$

For this system $CA = (0, 0, -1)$, but $CA^2 = (0, -1, 1)$. We can then show that this system is not SPR. To check this, by definition we take a pole in the complex right half-plane $s = 0.01 + 2i$, which results in

$$\text{Re}C(sI-A)^{-1}B = -0.0047. \quad (20)$$

Since the transfer function value is negative, the system is not positive real and thus it is not passive. Moreover, there exists an integral controller which makes this system unstable, for example one with the control vector $\gamma = 1$ (since only one node can be controlled) and the gain $\kappa = 3$ (although with $\kappa = 2$ the system is still stable). This confirms our understanding that

the novel $CA^2 > 0$ condition is meant to ensure stability using any arbitrary boundary node and arbitrary positive gain κ .

Going deeper to understand topological properties of the condition $CA^2 > 0$, we first start with more intuitive one, $CA < 0$, which is implied by $CA^2 > 0$.

Define matrices D and E such that $A = E - D$, with D being diagonal and E having all diagonal elements zero. Thus both D and E have all their entries positive. Matrix E can be viewed as adjacency matrix of the network, with element E_{ij} meaning influence of node v_j on node v_i . Matrix D consists of the self-damping powers on the diagonal. Therefore condition $CA < 0$ reads as $CD > CE$. This condition states some kind of diagonal dominance in the network.

Assume some $C_i = 0$. Then $(CD)_i = 0$ because D is diagonal. Thus $(CE)_i$ should be also zero, which means that for every index j either $C_j = 0$ or $E_{ji} = 0$.

Corollary 2. *If node v_i is not included in the aggregated output ($C_i = 0$), then its reachable set should not be included either.*

- For a strongly connected graph this means that all nodes should be included in the aggregated output.
- If the network is divided into "boundary" nodes and "inner" nodes, and the goal is to control the average of the inner nodes, then at least one of the boundary nodes should also be included into the average.

In the same manner it is possible to see this condition as a lower bound on the damping of each node: $D_{ii} \geq \sum_j C_j E_{ji} / C_i$. Thus the bigger is the influence of the node's neighbours in the output, the bigger should be the node's damping.

We can use the same decomposition $A = E - D$ in order to understand the condition $CA^2 > 0$ and conclude that

$$CE^2 + CD^2 > C(ED + DE). \quad (21)$$

Being a quadratic inequality, this condition bounds damping of each node from above and below with respect to dampings of other nodes. We will see examples in the following section.

E. Examples for average control

Here we present three examples of networks, namely a star, a line and an Erdős-Rényi graph, and analyse the condition $CA^2 > 0$ for them.

1) *Network with star topology:* To begin with we choose network with star topology with one central node and n leafs, average state of which we want to control. Let nodes $1 \dots n$ be the leafs and node $n+1$ be the center. Assume the center and the first leaf belong to the boundary node set and thus can be controlled (see Fig. 5). Dynamics of this network can be written as system (2) with matrices

$$A = \begin{pmatrix} -1-\beta & 0 & \dots & 1 \\ 0 & -1-\beta & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & -n-\alpha \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{pmatrix}, \quad (22)$$

$$C = \mathbf{1}^T / (n+1). \quad (23)$$

Such choice of system matrices corresponds to the undirected network with star topology and damping $\alpha > 0$ for central

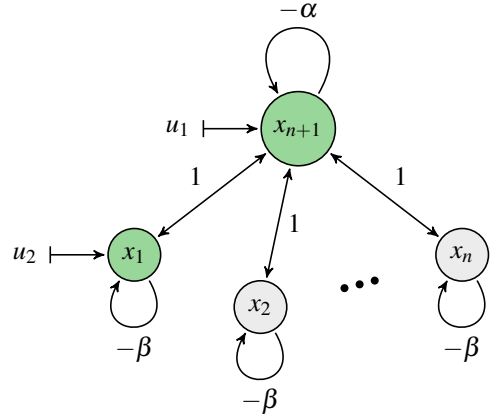


Figure 5. Network with star topology with n leafs. Boundary nodes in green.

node and $\beta > 0$ for all other nodes. The choice of B explores both cases of controlling leaf and center. It allows for maximum generality, moreover the controllability is guaranteed by Corollary 1. Integral controller (11) with $\gamma = (1, 0)^T$ would correspond to the control applied only to the center, and controller with $\gamma = (0, 1)^T$ would correspond to the control of the first leaf.

Calculating CA and CA^2 gives

$$\begin{aligned} CA &= (-\beta \quad -\beta \quad \dots \quad -\alpha) / (n+1) < 0, \\ (CA^2)_{1 \dots n} &= (\beta^2 + (\beta - \alpha)) / (n+1), \\ (CA^2)_{n+1} &= (\alpha^2 + n(\alpha - \beta)) / (n+1). \end{aligned} \quad (24)$$

$CA^2 > 0$ means then $\alpha^2 + n(\alpha - \beta) \geq 0$ and $\beta^2 + (\beta - \alpha) \geq 0$ with at least one of these inequalities being strict. Solving this for damping of leaf nodes we obtain

$$\sqrt{\alpha + \frac{1}{4}} - \frac{1}{2} \leq \beta \leq \alpha + \frac{\alpha^2}{n}, \quad (25)$$

thus β is bounded from both sides with respect to α . Moreover, as $n \rightarrow \infty$, we obtain limit inequality $\beta \leq \alpha$, which means that damping for leafs should be lower than damping for the center.

Simulation results for both cases, $\gamma = (1, 0)^T$ and $\gamma = (0, 1)^T$, and for $n = 20$ leafs are given on Fig. 6, with dampings $\alpha = 2$, $\beta = 1.1$, desired output value $y_d = 5$ and integral controller gain $\kappa = 12$. On Fig. 6a it is clearly seen that controlling the central node and controlling the leaf has almost the same effect on the output y .

2) *Line network:* Now we explore an example of a directed line network with n nodes. This network is depicted on Fig. 7. As usual, we are interested in controlling average state of the network, and it is assumed that we can control only the input node x_1 of the system. System matrices for n nodes are given as follows:

$$A = \begin{pmatrix} -1-\beta & 0 & \dots & 0 \\ 1 & -1-\beta & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -\alpha \end{pmatrix}, B = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (26)$$

$$C = \mathbf{1}^T / n, \quad \gamma = 1. \quad (27)$$

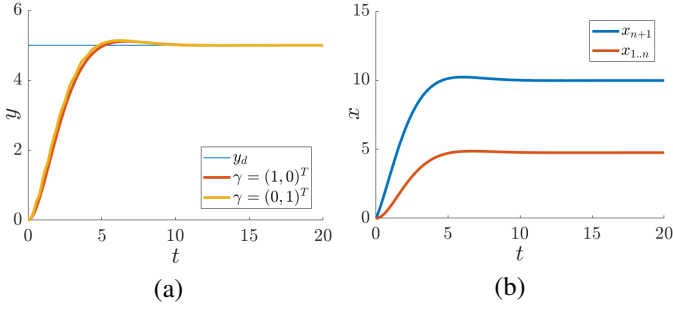


Figure 6. Output control of the star network with $n = 20$ leaves. $\alpha = 2$, $\beta = 1.1$, $\kappa = 12$, $y_d = 5$. (a). Output y for different γ vectors. (b). Spread of states x for $\gamma = (1, 0)^T$ corresponding to the control of the central node. All the leaf states $x_1..x_{20}$ have the same asymptotic value 4.751 (which is obvious from the symmetry), while the central state converges to 9.978.

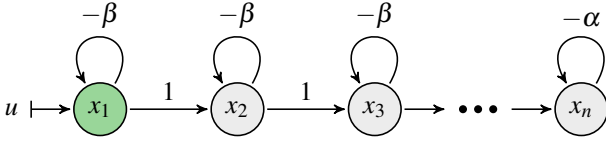


Figure 7. Directed line network with n nodes. Boundary node in green

This choice of system matrices corresponds to the directed line network with damping $\alpha > 0$ for the last node and $\beta > 0$ for all other nodes.

Calculating CA and CA^2 gives

$$\begin{aligned} CA &= (-\beta \quad -\beta \quad \dots \quad -\alpha)/n < 0, \\ (CA^2)_{1..n-2} &= \beta^2/n > 0, \\ (CA^2)_{n-1} &= (\beta^2 + \beta - \alpha)/n, \\ (CA^2)_n &= (\alpha^2 + \alpha - \beta)/n, \end{aligned} \quad (28)$$

$CA^2 > 0$ means then $\alpha^2 + (\alpha - \beta) \geq 0$ and $\beta^2 + (\beta - \alpha) \geq 0$. Solving this for damping of leaf nodes we obtain

$$\sqrt{\alpha + \frac{1}{4} - \frac{1}{2}} \leq \beta \leq \alpha + \alpha^2, \quad (29)$$

thus β is bounded from both sides with respect to α .

In order to validate the conclusions about this example, we take directed line networks with 4 and 100 nodes and check whether they are stable or unstable for different κ .

Fix $\alpha = 0.2$, therefore for condition $CA^2 > 0$ to hold one needs $\sqrt{0.45 - 0.5} \leq \beta \leq 0.24$. On Fig. 8 simulation results are shown for $\kappa = 12$, $y_d = 5$ and for two values of β , the first, $\beta = 0.2$, satisfies the condition, and the second $\beta = 0.002$ does not. In the case $\beta = 0.2$ and $n = 100$ it is very interesting to see what are limit values of the state variables x . It appears that they decrease exponentially starting from the controlled node x_1 , while preserving their average equal to y_d . This is due to the fact that in the steady state all node states except the first one and the last one should satisfy relation $x_{i-1} - (1 + \beta)x_i = 0$.

3) *Random Erdős-Rényi graph*: Here we present a simulation results for an integral controller for random Erdős-Rényi graph with $n = 4000$ nodes and probability of creating an edge $p = 0.01$. Vector $C = \mathbf{1}^T/n$ represents the average, and the

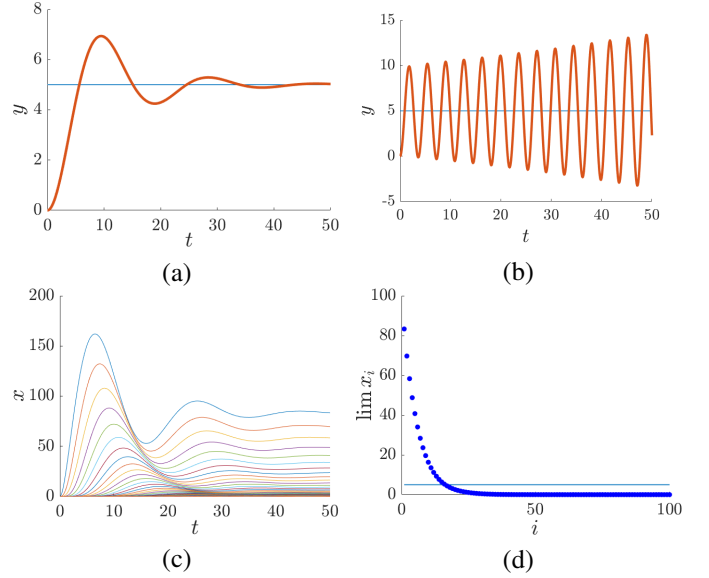


Figure 8. Output control of the directed line network, $\alpha = 0.2$, $\kappa = 12$, $y_d = 5$. Inset (b). $n = 4$, $\beta = 0.002$. Output y of the network is unstable, $CA^2 \not> 0$. Insets (a), (c), (d). $n = 100$, $\beta = 0.2$. Network is stable. (a). Output y . (c). Spread of states x . (d). Values of $\lim_{t \rightarrow \infty} x_i$ depending on the number $i \in \{1..100\}$, which is the distance from the controlled node.

system matrix A is the negative of a Laplacian of this ER graph with an additional random damping on every node taken from the uniform distribution $U(6, 7)$ such that $CA^2 > 0$. Matrix B is chosen to be a random vector of zeros and ones with equal probability. With such setup for any $\kappa > 0$ the system converges to the desired output reference $y_d = 5$, see Fig. 9.

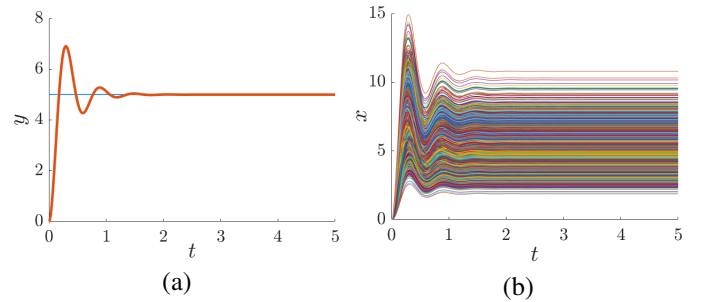


Figure 9. Output control of Erdős-Rényi graph for $n = 4000$ nodes, $\kappa = 250$. (a). Dynamics of output y . (b). Dynamics of states x .

IV. DEVIATION MINIMIZATION

A. Problem motivation

The previous sections showed that the control of average can be performed rather easily and without any knowledge of the system. But in some cases controlling the average with arbitrary control direction γ can lead to a poor performance: although the average state $y \rightarrow y_d$, states themselves can be very far from y_d . As an example one can look at Fig. 8c-d, where the spread of states x is shown. Although the average value is 5, most of the states in the steady state are almost zero, while some states are much larger, around 80.

This dispersion between states is captured by the squared deviation V . The smaller it is, the closer are the states to the average value.

Therefore, it makes sense to find the control law $u = u(y, V)$ for the system (2), which solves simultaneously two problems: assures $\lim_{t \rightarrow \infty} y(t) = y_d$ and minimizes the squared deviation V .

Preliminary, let us make the following observation. Controlling the average to the desired value in the steady state means that the system should satisfy one-dimensional constraint $-CA^{-1}Bu^* = y_d$, thus if the dimension of the steady-state control vector u^* is k , there are still $k-1$ degrees of freedom left for optimizing the control direction in sense of minimization of the squared deviation.

B. Explicit solution

Let us assume that the desired steady state is reached, and try to find it. Denote x^* and u^* as the state vector and the control vector respectively in the steady state. Also denote the steady-state squared deviation as V^* . Then the equations for the steady state, obtained from the system (2), assuming $y \rightarrow y_d$, are

$$\begin{cases} 0 = Ax^* + Bu^*, \\ y_d = Cx^*, \\ V^* = x^{*T}Px^*. \end{cases} \quad (30)$$

Our problem can be seen as a linear constrained quadratic minimization problem:

$$\begin{aligned} & \text{minimize } V^* = x^{*T}Px^*, \\ & \text{subject to } Ax^* + Bu^* = 0, \\ & Cx^* = y_d. \end{aligned} \quad (31)$$

In comparison to the standard linear-quadratic regulator, note that problem (31) is formulated for the steady state, thus there is no more dynamics in it, as well as no optimization of the transient process.

Assume for a moment that all the system matrices are known. Using the fact that the matrix A is stable, we can take the inverse and thus obtain the steady state vector $x^* = -A^{-1}Bu^*$. Denoting $S = B^T A^{-T} P A^{-1} B$ and $\eta = -B^T A^{-T} C^T$, we can write the minimization problem (31) in terms of u^* :

$$\begin{aligned} & \text{minimize } V^* = u^{*T} S u^*, \\ & \text{subject to } \eta^T u^* = y_d. \end{aligned} \quad (32)$$

Solution for the constrained problem is found using the Lagrangian:

$$L(u) = u^T S u + \lambda (\eta^T u - y_d). \quad (33)$$

Minimizing it over the control variable and solving for the Lagrange multiplier, we find that the explicit solution to the minimization problem is given by

$$\begin{cases} \lambda^* = -\frac{2y_d}{\eta^T S^{-1} \eta}, \\ u^* = \frac{y_d}{\eta^T S^{-1} \eta} S^{-1} \eta, \\ x^* = -\frac{y_d}{\eta^T S^{-1} \eta} A^{-1} B S^{-1} \eta. \end{cases} \quad (34)$$

Without loss of generality we will assume that S is positive definite for the future analysis.

The solution (34) cannot be used explicitly due to the fact that the system matrices are assumed to be unknown. But the next section introduces an algorithm which is able to stabilize the system in the arbitrary small neighbourhood of this solution.

C. Extremum seeking for squared deviation minimization

Extremum seeking is a form of adaptive control where the steady-state input-output characteristic is optimized, without requiring any explicit knowledge about this input-output characteristic other than that it exists and that it has an extremum [36], [37]. This algorithm, developed in the first part of XX century, explores the control space with small oscillations and provides an approximation of the gradient, which then can be integrated in order to find the optimum.

In standard realisations of extremum seeking, one adds to a current control input an oscillating signal, which should be small and slow in comparison with the system dynamics. Further, multiplying the output by the same oscillating signal, it is possible to recover an estimate of the gradient of the output with respect to the input.

This standard algorithm is unfortunately not usable for us, since we want to perform a constrained optimization (32) with a constraint that the average steady state should be equal to the desired one. However, if we modify the algorithm so to minimize the Lagrangian (33) instead of the squared deviation itself, we will optimize the original squared deviation while preserving the average state constraint. This modification leads to the introduction of Lagrange multiplier λ , which can be reconstructed by an additional integrator.

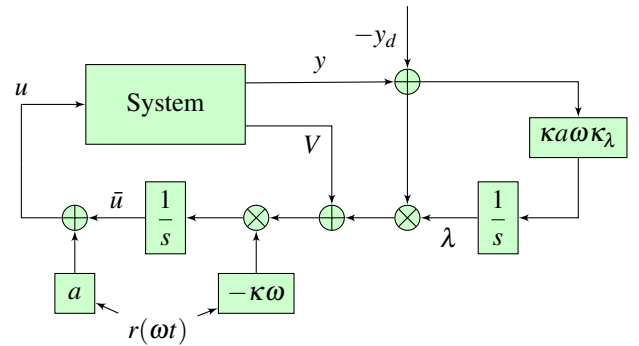


Figure 10. Extremum seeking scheme for constrained minimization

Assume the control law for the system (2) is given by

$$\begin{cases} \dot{u} = -\kappa\omega r(\omega t)(V + \lambda(y - y_d)), \\ \dot{\lambda} = \kappa\omega\kappa_\lambda(y - y_d), \\ u = \bar{u} + ar(\omega t), \end{cases} \quad (35)$$

where a and κ are small gains, ω is a small frequency, κ_λ is a relative Lagrange multiplier adaptation gain, and the oscillating signal $r(\omega t)$ is defined as

$$r(\omega t) = \sqrt{2} \cdot (\sin(2\pi\omega t) \quad \cos(2\pi\omega t) \quad \sin(4\pi\omega t) \quad \dots)^T.$$

Inputs to this control scheme (Fig. 10) are the output average y and output squared deviation V . Therefore, the control law does not use any state feedback or the system matrices. This is a multi-variable [41] extremum seeking control scheme, augmented with an additional integrator for the adaptation of the Lagrange multiplier as in primal-dual method [42], [43] with κ_λ being the relative speed of adaptation.

To begin with the proof of the stability of the control scheme (35), we first present a preliminary analysis, treating ω , κ and a as "small" parameters. Assuming ω is small, we define a new time-scale $\tau := \omega t$, which should be slow enough such that the dynamics of the system (2) is much faster than the dynamics of the adaptation. Under this time-scale the closed-loop system equations are

$$\begin{cases} \omega \frac{dx}{d\tau} = Ax + B(\bar{u} + ar(\tau)), \\ \frac{d\bar{u}}{d\tau} = -\kappa r(\tau) (x^T Px + \lambda(Cx - y_d)), \\ \frac{d\lambda}{d\tau} = \kappa a \kappa_\lambda (Cx - y_d). \end{cases} \quad (36)$$

With small ω the singular perturbation analysis [44] can be performed, thus system dynamics is substituted with its steady-state input-state mapping, i.e. $x^* = x^*(u) = -A^{-1}Bu$. The reduced dynamics is then approximated by

$$\begin{cases} \frac{d\bar{u}}{d\tau} = -\kappa r(\tau) [(\bar{u} + ar(\tau))^T S(\bar{u} + ar(\tau)) + \lambda(\eta^T(\bar{u} + ar(\tau)) - y_d)], \\ \frac{d\lambda}{d\tau} = \kappa a \kappa_\lambda (\eta^T(\bar{u} + ar(\tau)) - y_d). \end{cases} \quad (37)$$

As a next step we introduce an additional time-scale $\theta := \kappa \tau$, using which the system becomes

$$\begin{cases} \frac{d\bar{u}}{d\theta} = -r(\tau) [(\bar{u} + ar(\tau))^T S(\bar{u} + ar(\tau)) + \lambda \eta^T ((\bar{u} + ar(\tau)) - y_d)], \\ \frac{d\lambda}{d\theta} = a \kappa_\lambda (\eta^T(\bar{u} + ar(\tau)) - y_d). \end{cases} \quad (38)$$

This system is periodic in τ with unit period. When κ is small, the reduced dynamics can be approximated well by the dynamics averaged over the unit period:

$$\begin{cases} \frac{d\bar{u}_{av}}{d\theta} = - \int_0^1 \{ r(\sigma) [(\bar{u}_{av} + ar(\sigma))^T S(\bar{u}_{av} + ar(\sigma)) + \lambda_{av} (\eta^T(\bar{u}_{av} + ar(\sigma)) - y_d)] \} d\sigma, \\ \frac{d\lambda_{av}}{d\theta} = \int_0^1 \{ a \kappa_\lambda (\eta^T(\bar{u}_{av} + ar(\sigma)) - y_d) \} d\sigma. \end{cases} \quad (39)$$

Recall that by the definition of $r(\cdot)$ the oscillating signal has the following properties: $\int_0^1 r(\sigma) d\sigma = 0$ and $\int_0^1 r(\sigma) r(\sigma)^T d\sigma = I$. Then we can rewrite the system:

$$\begin{cases} \frac{d\bar{u}_{av}}{d\theta} = -2aS\bar{u}_{av} - a\eta\lambda_{av} - a^2R, \\ \frac{d\lambda_{av}}{d\theta} = a\kappa_\lambda \eta^T \bar{u}_{av} - a\kappa_\lambda y_d, \end{cases} \quad (40)$$

where $R = \int_0^1 r(\sigma) r(\sigma)^T S r(\sigma) d\sigma$. It can be seen that $2Su + \eta\lambda$ is the gradient of the Lagrangian, thus

$$\begin{cases} \frac{d\bar{u}_{av}}{d\tau} = -a\nabla_{\bar{u}_{av}} L + O(a^2), \\ \frac{d\lambda_{av}}{d\tau} = a\kappa_\lambda \nabla_{\lambda_{av}} L, \end{cases} \quad (41)$$

which converges to $O(a)$ of the explicit solution (u^*, λ^*) . Concretely, analysing steady state we obtain

$$\begin{aligned} y_{av}^* &= \eta^T \bar{u}_{av}^* \equiv y_d, \\ \lambda_{av}^* &= \lambda^* - a \frac{\eta^T S^{-1} R}{\eta^T S^{-1} \eta}, \\ \bar{u}_{av}^* &= u^* + \frac{a}{2} \left[\frac{\eta^T S^{-1} R}{\eta^T S^{-1} \eta} S^{-1} \eta - S^{-1} R \right]. \end{aligned} \quad (42)$$

The rigorous stability proof is based on the notion of semi-global practical asymptotic stability:

Definition 2 (SPA stability, [45]). Consider the parametrized family of systems:

$$\dot{\mathbf{x}} = f(t, \mathbf{x}, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_l), \quad (43)$$

where $\mathbf{x} \in \mathbb{R}^n$ and parameters of the system $\varepsilon_i > 0 \forall i = 1, 2, \dots, l$. The system (43) is said to be semi-globally practically asymptotically (SPA) stable in $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l]$ at \mathbf{x}^* , if there exists $\beta \in \mathcal{H}\mathcal{L}$ [46] such that the following holds: for each pair of strictly positive numbers (Δ, ν) , there exists $\varepsilon_1^* > 0$ and for any $\varepsilon_1 \in (0, \varepsilon_1^*)$ there exists $\varepsilon_2^* = \varepsilon_2^*(\varepsilon_1) > 0$ and for any $\varepsilon_2 \in (0, \varepsilon_2^*)$ there exists $\varepsilon_3^* = \varepsilon_3^*(\varepsilon_1, \varepsilon_2) > 0, \dots$, there exists $\varepsilon_l^* = \varepsilon_l^*(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{l-1}) > 0$ such that for any $\varepsilon_l \in (0, \varepsilon_l^*)$ the solutions of (43) with the parameters $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l]$ satisfy:

$$|\mathbf{x} - \mathbf{x}^*| \leq \beta(|\mathbf{x}_0 - \mathbf{x}^*|, (\varepsilon_1 \cdot \varepsilon_2 \cdots \varepsilon_l)(t - t_0)) + \nu \quad (44)$$

for all $t \geq t_0 \geq 0$, $\mathbf{x}(t_0) = \mathbf{x}_0$ with $|\mathbf{x}_0 - \mathbf{x}^*| \leq \Delta$.

Remark 1. Note that the order of the parameters $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l]$ is very important, because the bound for every parameter depends on the choice of all previous parameters, i.e. ε_3^* depends on the chosen ε_1 and ε_2 .

Theorem 5. System (2) with applied control law (35) is SPA stable in $[a, \kappa, \omega]$ at (x^*, u^*, λ^*) .

Proof. First we see that the system (40) is a linear system with the system matrix $M = \begin{pmatrix} -S & -\eta \\ \kappa_\lambda \eta^T & 0 \end{pmatrix}$, multiplied by a . The matrix M is stable, which can be shown by analysing its eigenvalues. Assuming $\mu \geq 0$ being eigenvalue of M , we show that the characteristic polynomial can have no roots. By the Schur complement:

$$\det(M - \mu I) = (-\kappa_\lambda \eta^T (S + \mu I)^{-1} \eta - \mu) \det(-S - \mu I). \quad (45)$$

Matrix S is positive definite, thus $S + \mu I$ is also positive definite. Therefore $\det(-S - \mu I) \neq 0$, $\eta^T (S + \mu I)^{-1} \eta > 0$ and $\mu \geq 0$, which means that (45) cannot equal zero.

We see that the matrix M have no non-negative eigenvalues, which means that the system (40) is SPA stable in a .

Then, using Lemma 1 from [45], we see that the system (37) is SPA stable in $[a, \kappa]$, and finally using Lemma 2 from

[45], we conclude that the original closed-loop system is SPA stable in $[a, \kappa, \omega]$, which concludes the proof. \square

Note that the stability of the closed-loop system (2)-(35) heavily depends on the chosen parameters. We proved SPA stability in $[a, \kappa, \omega]$, which by definition means that the bound for κ depends on the chosen α , and the bound for ω depends on the chosen α and κ . Therefore, it is difficult to find any rigorous bounds for how small these parameters should be. We can make only heuristic assumptions, such as requiring that all these parameters should be an order of magnitude smaller than the system impulse response.

Remark 2 (Additional integral controller). The extremum seeking scheme (35) can be enhanced by an array of possible modifications. For instance, an additional integral controller can be added:

$$\begin{cases} \dot{u} = -\kappa a \omega \gamma (y_d - y) - \kappa \omega r(\omega t) (V + \lambda (y - y_d)), \\ \dot{\lambda} = \kappa a \omega \kappa_\lambda (y - y_d), \\ u = \bar{u} + ar(\omega t), \end{cases} \quad (46)$$

where the vector γ such that $\eta^T \gamma > 0$ plays the same role as in the first sections of this paper regarding average control. I.e. for positive system it is enough to take $\gamma > 0$, as in Theorem 2.

Stability proof for the controller (46) follows exactly the same steps as in Theorem 5, replacing $S > 0$ by $S + \gamma \eta^T > 0$.

This scheme can provide much faster convergence of the average (see the examples), which means that the bigger adaptation gains can be used without the possibility for λ to diverge. The equilibrium point for the averaged reduced model of the closed-loop system with this scheme is (42), exactly the same as in the previous case.

There are a lot of other possible modifications of the extremum seeking scheme that can be usefully included, for example the high- and low-pass filters that are added to pick up the adaptation signal [37].

D. Examples for squared deviation minimization

All the algorithms presented before were tested on a graph constructed as a random Erdős-Rényi graph with $n = 40$, probability of creating an edge (with weight 1) $p = 0.1$ and self-loops with weight -5 . The dimension of the control vector was chosen $k = 3$, with matrix $B \in \mathbb{R}^{n \times k}$ being filled randomly: each element was set either to 0 or 1 with equal probability. Matrices C and P were chosen such that y corresponds to the average and V to the squared deviation of the states of the system. Desired value for the average was set $y_d = 5$.

In order to compare the speed of different algorithms we calculated characteristic times for the dynamics of average and squared deviation, defined as a negative inverse of the largest eigenvalue of the closed-loop system. Results of the simulations of different algorithms are presented in Figures 11-15.

To begin with, we apply the integral controller (11) to the system, and the dynamics of the average state y and the squared deviation V are shown on Fig. 11. The gain values are $\kappa = 1$ and $\gamma = [0, 0, 3]^T$. It is clearly seen that the squared deviation V does not reach its minimal value V^* , although this

controller is the fastest one: its characteristic time is $T = 3.84$.

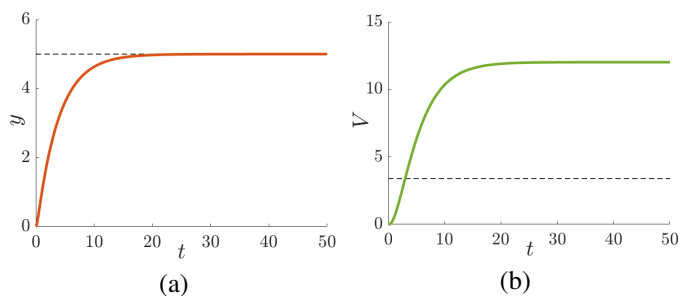


Figure 11. Integral controller (11) for average control, $T = 3.84$. (a). Average state y , with black dashed line denoting y_d . (b). Squared deviation V , with black dashed line denoting V^* .

Now we aim to minimize the deviation of the system states together with controlling the average. If the extremum seeking scheme (35) is used, the goal is achieved, and the performance is shown in Fig. 12. The gain values are $\omega = 0.1$, $\kappa = 2$, $a = 1$ and $\kappa_\lambda = 0.01$. This scheme is rather difficult to tune and also very slow, the characteristic time is $T = 2350$, three orders of magnitude higher than in Fig. 11. Also, significant oscillations in y , V and u can be seen even after the convergence of the system.

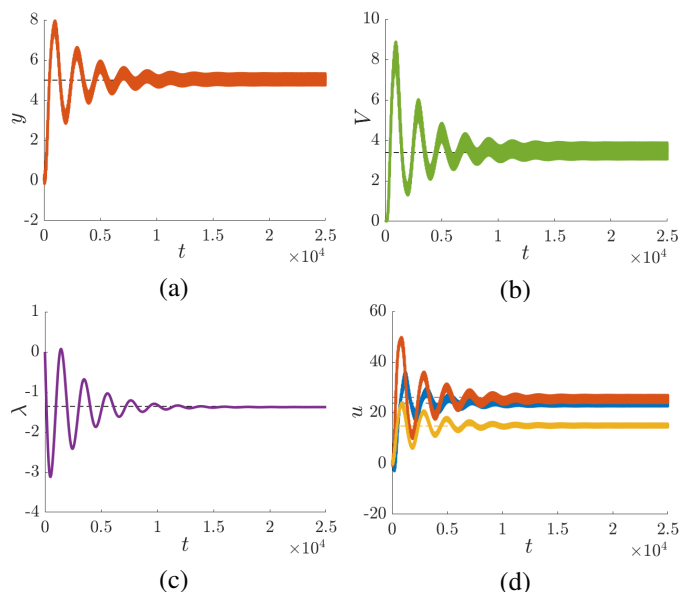


Figure 12. Extremum seeking control (35), $T = 2350$. (a). Average state y , with black dashed line denoting y_d . (b). Squared deviation V , with black dashed line denoting V^* . (c). Lagrange multiplier λ , with black dashed line denoting λ^* . (d). Control vector u , with dashed lines denoting u^* .

To minimize the oscillations in the extremum seeking, one needs to minimize the gains, but this would lead to an increased convergence time. Therefore we may try to improve the extremum seeking controller by adding the time-dependence to the gains, making them large at the beginning and decreasing them over time. In adaptation algorithms the gains should decrease slower than $1/t$ [47], otherwise the algorithm does not converge. In the extremum seeking (35) the gains a and κ are multiplied together, therefore their

product should decrease slower than $1/t$. In this example we set $a = \kappa = \frac{12}{(t/10+1)^{0.4}}$ together with $\omega = 0.1$ and $\kappa_\lambda = 0.01$. Performance of this scheme is shown in Fig. 13. It works much faster than the original one (the characteristic time is $T = 434$), and the oscillations are smaller, although the overshoot is larger.

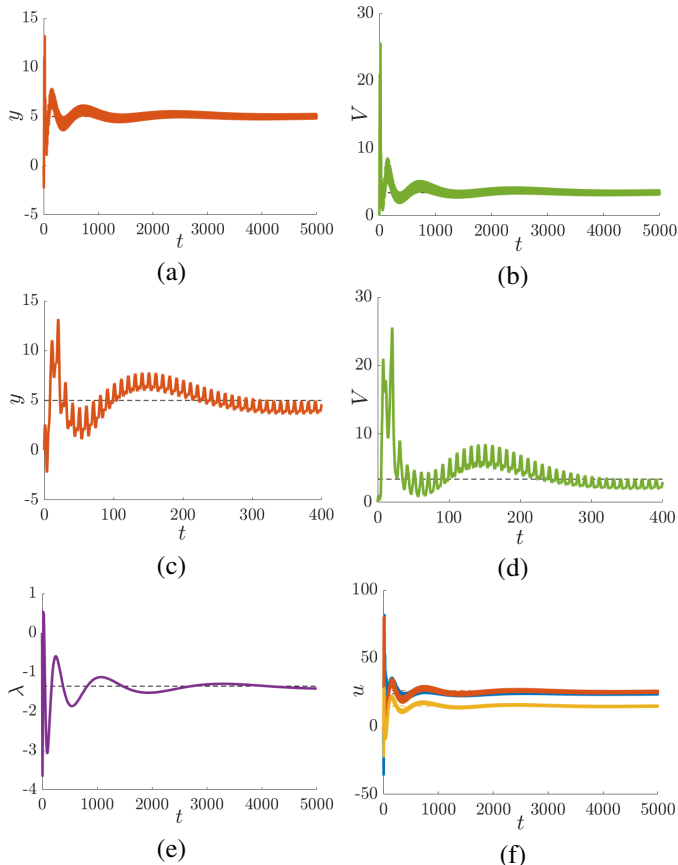


Figure 13. Extremum seeking control (35) with the gains decreasing over time, $T = 434$. (a). Average state y , with black dashed line denoting y_d . (b). Squared deviation V , with black dashed line denoting V^* . (c) and (d): Short-term plots for y and V . (e). Lagrange multiplier λ , with black dashed line denoting λ^* . (f). Control vector u , with dashed lines denoting u^* .

Performance of the extremum seeking scheme (46) is shown on Fig. 14. The gain values are $\omega = 0.15$, $\kappa = 2$, $a = 1$ and $\kappa_\lambda = 0.1$. With respect to the scheme (35), the gain κ_λ can be chosen larger. This leads to the faster adaptation of the Lagrange multiplier, thus this scheme works faster than (35). The dynamics for average and deviation now behave differently due to the additional controller for the average, therefore it makes sense to find separate characteristic times for them. The characteristic time for average is just $T_y = 3.9$, while for squared deviation it is $T_V = 732$. The parameter of the integral controller is the same as in the case of the integral controller for average, $\kappa\omega\gamma = [0, 0, 3]^T$.

Finally, the implementation of the scheme (46) with time-decreasing gains is presented in Fig. 15. The parameters are $\omega = 0.15$, $\kappa_\lambda = 0.1$, $\kappa\omega\gamma = [0, 0, 3]^T$, and the dependent gains are $a = \frac{3}{(t/100+1)^{0.4}}$ and $\kappa = \frac{7}{(t/100+1)^{0.4}}$. It is clearly seen that this scheme is much faster than all previous ones, with the

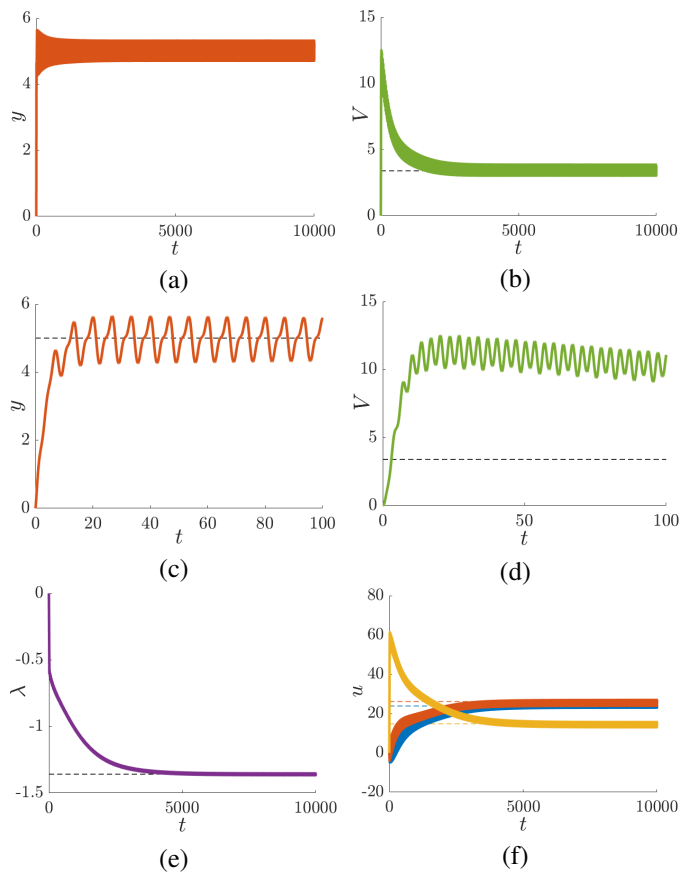


Figure 14. Extremum seeking control (46), $T_V = 732$ for squared deviation and $T_y = 3.9$ for average. (a). Average state y , with black dashed line denoting y_d . (b). Squared deviation V , with black dashed line denoting V^* . (c) and (d): Short-term plots for y and V . (e). Lagrange multiplier λ , with black dashed line denoting λ^* . (f). Control vector u , with dashed lines denoting u^* .

characteristic time for squared deviation being $T_V = 41$ and for average $T_y = 18$.

We see that the scheme (46) in general works faster than the scheme (35), but both of them are too slow to compare with the simple average controller (11). Their performance can be significantly increased using time-varying gains, although this leads to a large overshoot at the beginning.

V. CONCLUSION

In this paper, we considered the problem of the control of aggregates of a large-scale network system (namely, its average and variance). First we studied a linear output control problem and examined the general properties of the transfer functions of the system and the controller. We then studied the integral controller for the linear output regulation and formulated sufficient condition $CA^2 > 0$ for the convergence of any positive integral controller, giving in addition an example showing the conservatism of this condition. If the system satisfies this condition, the parameters of the controller can be chosen arbitrarily, and there is no need to have knowledge of the state vector or of the values of the elements of the A matrix. Control of the average state does not mean that the individual system states will be close to the average state. Therefore, in addition to controlling the average it is worth

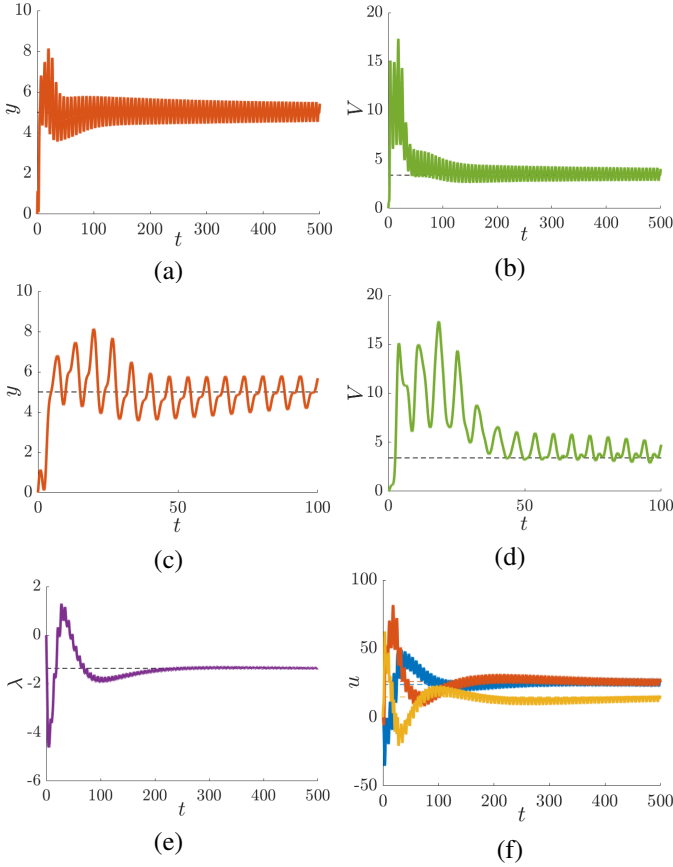


Figure 15. Extremum seeking control (46) with the gains decreasing over time, $T_V = 41$ and $T_y = 18$. (a). Average state y , with black dashed line denoting y_d . (b). Squared deviation V , with black dashed line denoting V^* . (c) and (d): Short-term plots for y and V . (e). Lagrange multiplier λ , with black dashed line denoting λ^* . (f). Control vector u , with dashed lines denoting u^* .

to minimize the deviation of the system states. To solve this problem we used the extremum seeking algorithm augmented with the primal-dual method for the constrained minimization. The stability of this scheme was proven and its performance, as well as that of several modified versions, was tested in the numerical simulations.

To conclude, we would like to further discuss the scope of application of this work. In the introduction, we argued that average and deviation can be directly measured in several practical examples. In an even broader range of cases, however, average and deviation can be estimated through sampling some nodes and constructing suitable observers, as recently illustrated in [27], [48]: the inclusion of such observers in our control scheme should be a topic of future work.

APPENDIX A PROOF OF THE LEMMA 1

By the definition of inverse $\mathcal{M}\mathcal{L} = (M + ibI)(L + i\bar{L}) = I$, which decomposes into real and imaginary parts:

$$\begin{aligned} ML - b\bar{L} &= I, \\ M\bar{L} + bL &= 0. \end{aligned} \quad (47)$$

From the second equation $\bar{L} = -bM^{-1}L$, and substitution of \bar{L} into the first equation gives

$$ML + b^2M^{-1}L = I, \quad (48)$$

which means $L = (M + b^2M^{-1})^{-1}$. \square

APPENDIX B PROOF OF THE LEMMA 2

Denote $L(t) = (M + tM^{-1})^{-1}$. We need to prove that $CL(t) > 0$ for all $t \geq 0$. The idea of the proof is to provide series expansion for $CL(t)$ such that each term in the expansion is positive. First the coefficients of Taylor series will be computed and then summation by parts will be used twice to obtain series with positive terms. The proof of the Lemma is separated into subsections A-F.

A. Series expansion

Matrix M is an M-matrix, which by definition means that there exists some matrix P with $P_{i,j} \geq 0$, $\rho(P) < 1$ and scalar $s > 0$ such that $M = s(I - P)$. Now make the following transformations:

$$\begin{aligned} L(t) &= (M + tM^{-1})^{-1} = M(M^2 + tI)^{-1} \\ &= s(I - P)(s^2(I - P)^2 + tI)^{-1} \\ &= s(I - P)((s^2 + t)I - 2s^2P + s^2P^2)^{-1} \\ &= \frac{s}{s^2 + t}(I - P)\left(I - \frac{2s^2}{s^2 + t}P + \frac{s^2}{s^2 + t}P^2\right)^{-1}. \end{aligned} \quad (49)$$

Multiplier $\frac{s}{s^2 + t}$ is always positive, thus it doesn't affect the sign of the result, so in future we will omit it. Now denote $\alpha = \frac{s^2}{s^2 + t}$. By definition of t and s this variable satisfies $0 < \alpha \leq 1$. Case $\alpha = 1$ is trivial (it corresponds to the case $t = 0$), thus often in the following we will use $0 < \alpha < 1$. Then

$$L = (I - P)(I - 2\alpha P + \alpha P^2)^{-1}. \quad (50)$$

We aim to find a coefficients in formal series expansion of L in the powers of P :

$$L = \sum_{k=0}^{+\infty} L_k P^k. \quad (51)$$

B. Coefficients of the series expansion

We can introduce scalar function $F(x)$ which has the same expansion as (51) and for which recursive computation of series coefficients is possible. Concretely, define

$$f(x) = \frac{1 - x}{1 - 2\alpha x + \alpha x^2}, \quad (52)$$

where $x \in [0, 1)$. Writing the same expansion as L :

$$f(x) = (1 - x) \sum_{k=0}^{+\infty} (2\alpha x - \alpha x^2)^k. \quad (53)$$

At the same time, $f(x)$ can be expanded as Taylor series centered at 0:

$$f(x) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!} x^k. \quad (54)$$

Power series expansion is unique, thus coefficients $L_k = \frac{f^{(k)}(0)}{k!}$.

The next step is to determine derivatives of $f(x)$ evaluated at $x = 0$. Let us introduce function

$$g(x) = 1 - 2\alpha x + \alpha x^2. \quad (55)$$

It is obvious that $f(x)g(x) = 1 - x$. Now take n -th derivative of this multiplication:

$$\frac{d^n}{dx^n} (1 - x) = \frac{d^n}{dx^n} (f(x)g(x)) = \sum_{k=0}^n \binom{n}{k} f^{(n-k)}(x)g^{(k)}(x). \quad (56)$$

Function $g(x)$ is a polynomial of the degree 2, thus its derivatives can be explicitly written:

$$g^{(0)}(0) = 1, \quad g^{(1)}(0) = -2\alpha, \quad g^{(2)}(0) = 2\alpha, \quad (57)$$

and all higher derivatives are zero. Moreover, $(1-x)^{(0)}(0) = 1$ and $(1-x)^{(1)}(0) = -1$ with all higher derivatives also zero. Recall that $L_n = \frac{f^{(n)}(0)}{n!}$. Using (56) we have the following recurrent relation for L_n :

$$L_n - 2\alpha L_{n-1} + \alpha L_{n-2} = 0, \quad \forall n \geq 2, \quad (58)$$

with initial conditions $L_0 = 1$ and $L_1 = 2\alpha - 1$.

C. Solving the linear recurrent equation

Equation (58) is a linear recurrent equation, which solution is found by solving the characteristic polynomial

$$\lambda^2 - 2\alpha\lambda + \alpha = 0. \quad (59)$$

For $0 < \alpha < 1$ roots are complex conjugate pair (λ, λ^*) with

$$\lambda = \alpha + i\sqrt{\alpha(1-\alpha)}, \quad |\lambda| = \sqrt{\alpha}. \quad (60)$$

The general solution to the equation (58) is given by $L_n = \text{Re}[z\lambda^n]$, where z is a complex value that should be determined from the initial conditions. From $L_0 = 1$ we simply recover $\text{Re}z = 1$, and from $L_1 = 2\alpha - 1$ it is found that $\text{Im}z = \sqrt{\frac{1-\alpha}{\alpha}}$. Thus the solution to the equation (58) is given by

$$L_n = \text{Re} \left[\left(1 + i\sqrt{\frac{1-\alpha}{\alpha}} \right) \left(\alpha + i\sqrt{\alpha(1-\alpha)} \right)^n \right]. \quad (61)$$

D. Back to the matrix equation

It is established that matrix L can be expressed by the series

$$L = \sum_{k=0}^{+\infty} L_k P^k, \quad (62)$$

where L_k are given by (61). Now it is evident that (62) is a convergent series due to $\rho(P) < 1$ and $|\lambda| = \sqrt{\alpha} < 1$.

Coefficients L_k can be both positive and negative, thus in general matrix L should not be positive. But we want to prove positivity of the vector CL :

$$CL = \sum_{k=0}^{+\infty} L_k CP^k > 0. \quad (63)$$

E. Properties of $\{CP^k\}$ sequence

Now it is time to use the condition $CM^2 > 0$. First of all, M is an M-matrix, thus for any vector x inequality $xM > 0$ implies $x > 0$. Therefore $CM > 0$ (and actually $C > 0$ automatically).

From $CM > 0$ we obtain $C(I-P) > 0$, which means $C > CP$. Moreover, matrix P is positive, thus multiplying both sides of this inequality on P preserves it. Thus the order relation holds:

$$C > CP > CP^2 > CP^3 > \dots > 0. \quad (64)$$

Therefore sequence $\{CP^k\}$ is monotonically decreasing with a limit zero (because $\rho(P) < 1$).

Now let us use the next condition, $CM^2 > 0$. Essentially it means $C(I-P)(I-P) > 0$, or $(C-CP) > (CP-CP^2)$. Again, multiplication by P preserves order, so we have

$$C - CP > CP - CP^2 > CP^2 - CP^3 > \dots > 0, \quad (65)$$

or

$$CP^k - 2CP^{k+1} + CP^{k+2} > 0. \quad (66)$$

This implies that sequence $\{CP^k - CP^{k+1}\}$ is also monotonically decreasing to zero. In some sense this is equivalent to the "convexity" of $\{CP^k\}$ sequence.

F. Summations by parts

For any series

$$\sum_{k=0}^N x_k y_k = x_N Y_N - \sum_{k=0}^{N-1} (x_{k+1} - x_k) Y_k, \quad (67)$$

where $Y_n = \sum_{k=0}^n y_k$. This transformation is called Abel transformation or summation by parts. We will apply this procedure twice to obtain series with each term positive.

Denote $H_n = \sum_{k=0}^n L_k$. Then H_n is bounded because L_k consists of powers of λ with $|\lambda| = \sqrt{\alpha} < 1$. By $\rho(P) < 1$ follows $\lim_{k \rightarrow +\infty} CP^k = 0$. Thus $\lim_{k \rightarrow +\infty} H_k CP^k = 0$ and we can write

$$\begin{aligned} CL &= \sum_{k=0}^{+\infty} L_k CP^k = - \sum_{k=0}^{+\infty} H_k (CP^{k+1} - CP^k) = \\ &= \sum_{k=0}^{+\infty} H_k (CP^k - CP^{k+1}). \end{aligned} \quad (68)$$

Applying Abel transformation for the second time with $G_n = \sum_{k=0}^n H_k$, we get

$$CL = \sum_{k=0}^{+\infty} G_k (CP^k - 2CP^{k+1} + CP^{k+2}). \quad (69)$$

Let us calculate H_n :

$$H_n = \sum_{k=0}^n L_k = \text{Re} \left[z \sum_{k=0}^n \lambda^k \right] = \text{Re} \left[z \frac{1 - \lambda^{n+1}}{1 - \lambda} \right], \quad (70)$$

where $\lambda = \alpha + i\sqrt{\alpha(1-\alpha)}$ and $z = 1 + i\sqrt{\frac{1-\alpha}{\alpha}}$. Multiply nominator and denominator by $(1 - \lambda^*)$:

$$H_n = \frac{1}{1 - \alpha} \text{Re} [z(1 - \lambda^*)(1 - \lambda^{n+1})]. \quad (71)$$

Product $z(1 - \lambda^*) = i\sqrt{\frac{1-\alpha}{\alpha}}$, which is purely imaginary, so

$$H_n = -\operatorname{Re} \left[i \frac{\lambda}{\sqrt{\alpha(1-\alpha)}} \lambda^n \right]. \quad (72)$$

Denote $w = -i\frac{\lambda}{\sqrt{\alpha(1-\alpha)}}$ and calculate G_n :

$$G_n = \sum_{k=0}^n H_k = \operatorname{Re} \left[w \sum_{k=0}^n \lambda^k \right] = \operatorname{Re} \left[w \frac{1 - \lambda^{n+1}}{1 - \lambda} \right]. \quad (73)$$

Multiply nominator and denominator by $(1 - \lambda^*)$:

$$G_n = \frac{1}{1 - \alpha} \operatorname{Re} [w(1 - \lambda^*)(1 - \lambda^{n+1})]. \quad (74)$$

Product $w(1 - \lambda^*) = 1$, thus this function reads as

$$G_n = \frac{1}{1 - \alpha} (1 - \operatorname{Re} [\lambda^{n+1}]). \quad (75)$$

By definition $|\lambda| = \sqrt{\alpha} < 1$, thus $\operatorname{Re} [\lambda^n] < 1$ for any $n > 0$. This means that $G_n > 0$ for any $n \geq 0$. Furthermore, by convexity of the sequence $\{CP^k\}$ for any $k \geq 0$: $CP^k - 2CP^{k+1} + CP^{k+2} > 0$.

Thus every term in (69) is greater than zero, which concludes the proof. \square

APPENDIX C PROOF OF THE LEMMA 3

As in previous proof, define $L = (M + tM^{-1})^{-1}$ and $M = s(I - P)$. By Lemma 2 $CL > 0$. Using the series expansion (69), we can write

$$CL = \sum_{k=0}^{+\infty} G_k (C - 2CP + CP^2) P^k > 0, \quad (76)$$

where all $G_k > 0$. Condition $CM^2B\gamma > 0$ reads as

$$(C - 2CP + CP^2)B\gamma > 0. \quad (77)$$

Then

$$\begin{aligned} CLB\gamma &= \sum_{k=0}^{+\infty} G_k (C - 2CP + CP^2) P^k B\gamma = \\ &= G_0 (C - 2CP + CP^2) B\gamma + \\ &+ \sum_{k=1}^{+\infty} G_k (C - 2CP + CP^2) P^k B\gamma, \end{aligned} \quad (78)$$

where the first term is strictly greater than zero and all others a greater or equal than zero. Thus $CLB\gamma > 0$, which concludes the proof. \square

ACKNOWLEDGMENT

The Scale-FreeBack project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement N 694209).

REFERENCES

- [1] G. Bianchin and F. Pasqualetti, "Gramian-based optimization for the analysis and control of traffic networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 3013–3024, 2019.
- [2] A. Mirtabatabaei and F. Bullo, "Opinion dynamics in heterogeneous networks: Convergence conjectures and theorems," *SIAM Journal on Control and Optimization*, vol. 50, no. 5, pp. 2763–2785, 2012.
- [3] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on automatic control*, vol. 51, no. 3, pp. 401–420, 2006.
- [4] F. Dorfler and F. Bullo, "Synchronization and transient stability in power networks and nonuniform kuramoto oscillators," *SIAM Journal on Control and Optimization*, vol. 50, no. 3, pp. 1616–1642, 2012.
- [5] G. Yan, J. Ren, Y.-C. Lai, C.-H. Lai, and B. Li, "Controlling complex networks: How much energy is needed?" *Physical review letters*, vol. 108, no. 21, p. 218703, 2012.
- [6] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *nature*, vol. 473, no. 7346, pp. 167–173, 2011.
- [7] N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, and C. T. Bergstrom, "Nodal dynamics, not degree distributions, determine the structural controllability of complex networks," *PLoS one*, vol. 7, no. 6, p. e38398, 2012.
- [8] N. Bof, G. Baggio, and S. Zampieri, "On the role of network centrality in the controllability of complex networks," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 643–653, 2016.
- [9] G. Lindmark and C. Altafini, "A driver node selection strategy for minimizing the control energy in complex networks," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8309–8314, 2017.
- [10] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," *Nature Physics*, vol. 11, no. 9, pp. 779–786, 2015.
- [11] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.
- [12] L. Vassio, F. Fagnani, P. Frasca, and A. Ozdaglar, "Message passing optimization of harmonic influence centrality," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 109–120, 2014.
- [13] W. S. Rossi and P. Frasca, "On the convergence of message passing computation of harmonic influence in social networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 2, pp. 116–129, 2018.
- [14] I. Klickstein, A. Shirin, and F. Sorrentino, "Energy scaling of targeted optimal control of complex networks," *Nature communications*, vol. 8, no. 1, pp. 1–10, 2017.
- [15] P. Wittmuess, M. Heidingsfeld, and O. Sawodny, "Optimal actuator design for optimal output controllability," *IFAC-PapersOnLine*, vol. 49, no. 21, pp. 234–239, 2016.
- [16] I. Klickstein, A. Shirin, and F. Sorrentino, "Locally optimal control of complex networks," *Physical review letters*, vol. 119, no. 26, p. 268301, 2017.
- [17] C. Commault, J. van der Woude, and P. Frasca, "Functional target controllability of networks: structural properties and efficient algorithms," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1521–1530, 2019.
- [18] G. Casadei, C. Canudas-de Wit, and S. Zampieri, "Controllability of large-scale networks: An output controllability approach," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 5886–5891.
- [19] M. Eslami and K. Faez, "Automatic traffic monitoring from satellite images using artificial immune system," in *Structural, Syntactic, and Statistical Pattern Recognition*, E. R. Hancock, R. C. Wilson, T. Windeatt, I. Ulusoy, and F. Escolano, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 170–179.
- [20] G. Palubinskas, F. Kurz, and P. Reinartz, "Model based traffic congestion detection in optical remote sensing imagery," *European Transport Research Review*, vol. 2, no. 2, pp. 85–92, Jun 2010.
- [21] A. Boynard *et al.*, "First simultaneous space measurements of atmospheric pollutants in the boundary layer from iasi: A case study in the north china plain," *Geophysical Research Letters*, vol. 41, no. 2, pp. 645–651, 2014.
- [22] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Transactions on biomedical engineering*, vol. 44, no. 9, pp. 867–880, 1997.

- [23] A. Hunt, I. Rusli, M. Schakel, and A. Kenbar, "High-speed density measurement for lng and other cryogenic fluids using electrical capacitance tomography," *Cryogenics*, vol. 113, p. 103207, 2021.
- [24] C. Canudas-de Wit, "Erc: Scale-free control for complex physical network systems," <http://scale-freeback.eu>, 2015.
- [25] G. Casadei, C. Canudas-de-Wit, and S. Zampieri, "Model reduction based approximation of the output controllability gramian in large-scale networks," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 4, pp. 1778–1788, 2020.
- [26] M. U. B. Niazi, C. Canudas-de Wit, and A. Y. Kibangou, "Average observability of large-scale network systems," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 1506–1511.
- [27] —, "Average state estimation in large-scale clustered network systems," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 4, pp. 1736–1745, 2020.
- [28] N. Martin, P. Frasca, and C. Canudas-de Wit, "Large-scale network reduction towards scale-free structure," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 4, pp. 711–723, 2019.
- [29] N. Martin, P. Frasca, and C. Canudas-de-Wit, "Subgraph detection for average detectability of lti systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2787–2798, 2020.
- [30] A. Rantzer, "Distributed control of positive systems," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 6608–6611.
- [31] —, "Scalable control of positive systems," *European Journal of Control*, vol. 24, pp. 72–80, 2015.
- [32] C. Briat, "Robust stability and stabilization of uncertain linear positive systems via integral linear constraints: L1-gain and ∞ -gain characterization," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 17, pp. 1932–1954, 2013.
- [33] F. N. Nogueira, "Contributions to positive systems control: mathematical methods and applications," Ph.D. dissertation, Universidade do Porto (Portugal), 2013.
- [34] R. Sepulchre, M. Jankovic, and P. V. Kokotovic, *Constructive nonlinear control*. Springer Science & Business Media, 2012.
- [35] N. Kottenstette, M. J. McCourt, M. Xia, V. Gupta, and P. J. Antsaklis, "On relationships among passivity, positive realness, and dissipativity in linear systems," *Automatica*, vol. 50, no. 4, pp. 1003–1016, 2014.
- [36] K. B. Ariyur and M. Krstić, *Real time optimization by extremum seeking control*. Wiley Online Library, 2003.
- [37] Y. Tan, W. H. Moase, C. Manzie, D. Nešić, and I. M. Mareels, "Extremum seeking from 1922 to 2010," in *Proceedings of the 29th Chinese control conference*. IEEE, 2010, pp. 14–26.
- [38] D. Nikitin, C. Canudas-de Wit, and P. Frasca, "Boundary control for output regulation in scale-free positive networks," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 5050–5055.
- [39] R. J. Plemmons, "M-matrix characterizations. i—nonsingular m-matrices," *Linear Algebra and its Applications*, vol. 18, no. 2, pp. 175–188, 1977.
- [40] K. Fan, "Topological proofs for certain theorems on matrices with non-negative elements," *Monatshefte für Mathematik*, vol. 62, no. 3, pp. 219–237, 1958.
- [41] W. H. Moase, Y. Tan, D. Nešić, and C. Manzie, "Non-local stability of a multi-variable extremum-seeking scheme," in *2011 Australian Control Conference*. IEEE, 2011, pp. 38–43.
- [42] A. Nedić and A. Ozdaglar, "Subgradient methods for saddle-point problems," *Journal of optimization theory and applications*, vol. 142, no. 1, pp. 205–228, 2009.
- [43] J. W. Simpson-Porco, B. K. Poolla, N. Monshizadeh, and F. Dörfler, "Input–output performance of linear–quadratic saddle-point algorithms with application to distributed resource allocation problems," *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2032–2045, 2019.
- [44] P. Kokotović, H. K. Khalil, and J. O'reilly, *Singular perturbation methods in control: analysis and design*. SIAM, 1999.
- [45] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.
- [46] H. K. Khalil and J. W. Grizzle, *Nonlinear systems*. Prentice hall Upper Saddle River, NJ, 2002, vol. 3.
- [47] V. S. Borkar, *Stochastic approximation: a dynamical systems viewpoint*. Springer, 2009, vol. 48.
- [48] M. U. B. Niazi, C. Canudas-de-Wit, and A. Y. Kibangou, "State variance estimation in large-scale network systems," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 6052–6057.



Denis Nikitin received both the B.Sc. and the M.Sc. degree in Mathematics and Mechanics Faculty of Saint Petersburg State University, Saint Petersburg, Russia, specializing on the control theory and cybernetics. He won several international robotics competitions while being student and was a teacher of robotics in the Math&Phys Lyceum 239 in Saint Petersburg. He is currently a doctoral researcher at CNRS, GIPSA-Lab, Grenoble, France. His research mainly focuses on control of large-scale network aggregates.



Carlos Canudas-de-Wit (F'16) was born in Villahermosa, Mexico, in 1958. He received the B.S. degree in electronics and communications from the Monterrey Institute of Technology and Higher Education, Monterrey, Mexico, in 1980, and the M.S. and Ph.D. degrees in automatic control from the Department of Automatic Control, Grenoble Institute of Technology, Grenoble, France, in 1984 and 1987, respectively. He is currently a Directeur de recherche (Senior Researcher) with CNRS, Grenoble, where he is the Leader of the NeCS Team, a joint team of GIPSA-Lab (CNRS) and INRIA, on networked controlled systems.

Dr. Canudas-de-Wit is an IFAC Fellow. He was an Associate Editor of the IEEE Transactions on Automatic Control, the Automatica, the IEEE Transactions on Control Systems Technology. He is an Associate Editor of the Asian Journal of Control, and the IEEE Transactions on Control of Network Systems. He served as the President of the European Control Association from 2013 to 2015, and a member of the IEEE Board of Governors of the Control System Society from 2011 to 2014. He holds the ERC Advanced Grant Scale-FreeBack from 2016 to 2021.



Paolo Frasca (M'13–SM'18) received the Ph.D. degree from Politecnico di Torino, Turin, Italy, in 2009. After Postdoctoral appointments with the CNR-IAC, Rome, and in Torino, he has been an Assistant Professor with the University of Twente, Enschede, The Netherlands, from 2013 to 2016. Since October 2016, he has been a CNRS Researcher with GIPSA-lab, Grenoble, France.

His research interests focus on the theory of network systems and cyber-physical systems, with applications to robotics, sensors, infrastructural, and social networks. On these topics, he has (co)authored more than 60 journal and conference papers and the book *Introduction to Averaging Dynamics Over Networks* (Springer). He is or has been an Associate Editor for the Editorial Boards of several conferences and journals, including the International Journal of Robust and Nonlinear Control and the IEEE Control Systems Letters.