Network-decentralised optimisation and control
An explicit saturated solution

Blanchini, Franco; Casagrande, Daniele; Fabiani, Filippo; Giordano, Giulia; Pesenti, Raffaele

DOI
10.1016/j.automatica.2019.02.009

Publication date
2019

Document Version
Accepted author manuscript

Published in
Automatica

Network-decentralised optimisation and control: an explicit saturated solution

Franco Blanchini [a], Daniele Casagrande [b], Filippo Fabiani [c], Giulia Giordano [d], Raffaele Pesenti [e]

[a]*Dipartimento di Matematica, Informatica e Fisica, University of Udine, Via delle Scienze, 206 - 33100 Udine, Italy (e-mail: blanchini@uniud.it)*

[b]*Dipartimento Politecnico di Ingegneria e Architettura, University of Udine, Via delle Scienze, 208 - 33100 Udine, Italy (e-mail: daniele.casagrande@uniud.it)*

[c]*Department of Information Engineering, University of Pisa, Via G. Caruso, 16 - 56122 Pisa, Italy (e-mail: filippo.fabiani@ing.unipi.it)*

[d]*Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2 - 2628 CD Delft, The Netherlands (e-mail: g.giordano@tudelft.nl)*

[e]*Dipartimento di Management, Università Ca' Foscari Venezia - Cannaregio 873 - 30121 Venezia, Italy (e-mail: pesenti@unive.it)*

Abstract

This paper proposes a decentralised explicit (closed-form) iterative formula that solves convex programming problems with linear equality constraints and interval bounds on the decision variables. In particular, we consider a team of decision agents, each setting the value of a subset of the variables, and a team of information agents, in charge of ensuring that the equality constraints are fulfilled. The structure of the constraint matrix imposes a communication pattern between decision and information agents, which can be represented as a bipartite graph. We associate each information agent with an integral variable and each decision agent with a saturated function, which takes the interval bounds into account, and we design a decentralised dynamic mechanism that globally converges to the optimal solution. Under mild conditions, the convergence is shown to be exponential. We also provide a discrete-time algorithm, based on the Euler system, and we give an upper bound for the step parameter to ensure convergence. Although the considered optimisation problem is static, we show that the proposed scheme can be successfully applied to find the optimal solution of network-decentralised dynamic control problems.

1 Introduction and Motivation

The complexity of large-scale systems, such as power distribution systems, water systems distributed in space, logistics and transportation systems, leads to severe difficulties in the definition of centralized optimisation and control policies, which typically arise due to systems' dimensionality, information structure constraints, uncertainty, and delays, and motivate the interest in decentralized optimisation and control (see, e.g., Bakule 2008, 2014). In recent years, this interest is possibly increased given the necessity of making strategic large-scale systems resilient to failures and attacks.

Distributed and decentralised optimisation and control traces back to the 60s. Sekine (1963), for example, proposed a decentralised optimisation setup in which the cost functional is the sum of local functionals depending on local variables, each associated with a decision agent, and a single coupling linear constraint is considered. This scheme was adopted in many subsequent contributions (see, for instance, Chang, Nedić & Scaglione 2014; Kozma, Conte & Diehl 2015; Chatzipanagiotis & Zavlanos 2016 and the references therein). Cherukuri & Cortés (2016) proposed a slightly different setup that considers a set of coupling linear constraints. In particular, the *constraint (coefficient) matrix* is assumed block-diagonal and the proposed algorithm exploits its structure. Other contributions consider agents that reach an agreement on the decision variables (Terelius, Topcu & Murray 2011; Falsone et al. 2017; Lin, Ren & Farrell 2017; Mokhtari, Ling & Ribeiro 2017): equality constraints are introduced for the decision variables associated with the agents. Several techniques were proposed to solve optimisation problems within the mentioned schemes: alternating direction method (Boyd et al. 2011; Makhdoumi & Ozdaglar 2017), nonuniform gradient gains (Lin, Ren & Farrell 2017), sub-gradient methods (Johansson, Rabi & Johansson 2009; Nedić & Ozdaglar 2009), dual averaging (Duchi, Agarwal & Wainwright 2002; Nedić, Lee & Raginsky 2015), extremum seeking control (Dougherty & Guay 2017). Bagagiolo et al. (2017) recently introduced a mean field game approach to routing problems for interacting agents distributed on a net-

Preprint submitted to Automatica *March 6, 2019*

work and gave an explicit expression of an optimal decentralised control via a state-space extension technique.

In this paper, we consider a convex programming problem with linear equality constraints, and upper and lower bounds on the decision variables. The optimisation problem, and in particular the connection between agents, is seen from a novel perspective that generalises previous schemes. The set of constraints, in fact, is generic and not related to any kind of predefined "coupling" between the decision variables, while the sparsity structure of the constraint matrix (a block matrix) imposes a communication pattern between two kinds of agents: decision and information agents. A decision agent is associated with each column (or block-column) of the constraint matrix, while an information agent is associated with each row (or block-row) of the matrix. Decision agents choose the value of their subsets of free variables, under proper "bounding-box" constraints, while information agents ensure that the linear equality constraints are satisfied. Each decision agent can communicate only with the subset of the information agents corresponding to non-zero blocks in its block-column. It cannot communicate with the other decision agents. Conversely, each information agent is aware only of the actions of the decision agents that correspond to non-zero blocks in its block-row. It cannot communicate with the other information agents.

The proposed scheme is embedded within the theoretical framework of system dynamics: we provide an explicit (closed-form) expression of the iterative formula to get the solution, in a feedback form. In particular, we design a dynamic decentralised mechanism, described by an explicit formula, which asymptotically solves the problem of minimizing the cost $\sum_i f_i(u_i)$, where f_i are convex functions, in the presence of lower and upper bounds on u_i and under the linear constraint $Bu = w$, where B is a block-structured matrix. The key features of the scheme are summarised next.

- Each information agent is equipped with an integral variable, expressing the violation of a subset of the constraints. The convergence of this integral variable ensures that the corresponding constraints are asymptotically satisfied; its dynamic equations only depend on the actions of the decision agents to which the information agent is connected.
- Each decision agent chooses its current action based on the integral variables of the information agents to which it is connected. An explicit, closed-form iterative formula given by a saturation function is provided.
- The dynamical solution scheme works in continuous time and its convergence to the optimal solution of the optimisation problem is proven by exploiting recent results from network-decentralised control (Bauso et al. 2013; Blanchini et al. 2016).
- From the continuous-time scheme, an iterative discrete-time algorithm is obtained by adopting the Euler discretisation method with step parameter $\tau > 0$. Convergence of the algorithm is ensured if $0 < \tau < \tau_{max} \propto 2/\|B\|^2$, where $\|B\|$ denotes the Euclidean norm of matrix B.
- Under mild conditions, the algorithm converges exponentially (i.e., "linearly" in the optimisation jargon).
- When the constraint matrix B is uncertain, a reduction of the step parameter bound τ_{max} is typically required.
- Although the scheme is thought for static problems, it is successfully applied to network-decentralised control problems (Iftar 1999; Iftar & Davison 2002; Blanchini, Franco & Giordano 2013, 2015), to derive optimal control strategies that are inherently decentralised.

2 Main Assumptions and Problem Statement

We consider the constrained optimisation problem

$$\begin{aligned} \min_{u \in \mathscr{U}} \quad & \sum_{i=1}^{m} f_i(u_i) \\ \text{s.t.} \quad & Bu = w\,, \end{aligned} \tag{1}$$

where functions $f_i : \mathscr{U} \to \mathbb{R}$, matrix $B \in \mathbb{R}^{n \times m}$, with $m > n$, and vector $w \in \mathbb{R}^n$ are given, and where the feasible set is

$$\mathscr{U} = \{u \in \mathbb{R}^m : u_i^- \leq u_i \leq u_i^+\,,\ \text{for } i = 1, \ldots, m\}\,, \tag{2}$$

for assigned lower and upper bound vectors u^- and u^+. To find the solution u^* of (1)–(2), we look for a decentralised dynamic mechanism that generates a function $u(t)$ such that

$$\lim_{t \to \infty} u(t) = u^*\,. \tag{3}$$

We call *decentralised solution* any solution of the optimisation problem (1)–(2) achieved in a decentralised manner. Throughout the paper, we denote by int$\mathscr{U}$ the interior of $\mathscr{U}$ and we make the following, standard assumptions.

Assumption 1 *For all $i = 1, \ldots, m$, the function f_i is strictly convex. In particular, it is twice continuously differentiable and there exists a constant $\mu > 0$ s.t., $\forall \chi$, $\frac{d^2}{d\chi^2} f_i(\chi) \geq \mu$.*

Assumption 2 *Matrix $B \in \mathbb{R}^{n \times m}$ is full row rank.*

Assumption 3 (Slater's constraint qualification) *There exists $u \in int\mathscr{U}$ for which $Bu = w$.*

Assumption 2 is necessary to have a proper optimisation problem, where the constraints $Bu = w$ define a non-empty subspace and none of them is redundant. Assumption 3 (cf. Boyd & Vandenberghe 2004) guarantees that the problem is strictly feasible, so that infinitesimal perturbations on w cannot lead to unfeasibility. However, in general the optimal solution u^* may not be in *int*$\mathscr{U}$.

We aim at solving problem (1)–(2) in a decentralised manner, by considering the interplay between a set $\mathscr{M}$ of *decision agents* and a set $\mathscr{N}$ of *information agents*. Decision and information agents communicate according to a bipartite graph corresponding to the block structure of the sparse matrix B (see, e.g., the graph in Fig. 1, associated with the constraint matrix in Example 2.1). Each decision agent $j \in \mathscr{M}$ iteratively fixes the tentative values of the m_j elements of a sub-vector u_j of u and each information agent $h \in \mathscr{N}$ verifies whether a subset $\mathscr{C}_h$ of n_h equality constraints of the system $\mathscr{C} : Bu = w$ are feasible for the values fixed by the decisions agents. More formally,

- each decision agent $j \in \mathscr{M}$ is associated with a subset of m_j consecutive indices in the set $\mathscr{I}_m = \{1, \ldots, m\}$; the

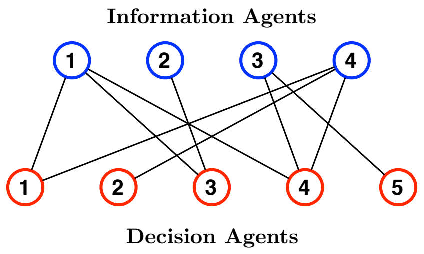

Figure 1. The information-decision graph for Example 2.1.

agents in $\mathscr{M}$ partition $\mathscr{I}_m$, so that $\sum_{j\in\mathscr{M}} m_j = m$;
- each information agent $h \in \mathscr{N}$ is associated with a subset of n_h consecutive indices in the set $\mathscr{I}_n = \{1,\dots,n\}$; the agents in $\mathscr{N}$ partition $\mathscr{I}_n$, so that $\sum_{h\in\mathscr{N}} n_h = n$.

Accordingly,

- vector u is partitioned into $|\mathscr{M}|$ sub-vectors u_j of dimension m_j, for $j \in \mathscr{M}$;
- the constraint set $\mathscr{C}$ is partitioned into $|\mathscr{N}|$ subsets $\mathscr{C}_h$ of cardinality n_h, while vector w is partitioned into $|\mathscr{N}|$ sub-vectors w_h of dimension n_h, for $h \in \mathscr{N}$;
- matrix B is partitioned into block-columns $B_{(:,j)} \in \mathbb{R}^{n\times m_j}$, each associated with a decision agent $j \in \mathscr{M}$, and into block-rows $B_{(h,:)} \in \mathbb{R}^{n_h\times m}$, each associated with an information agent $h \in \mathscr{N}$; these partitions identify $|\mathscr{N}| \times |\mathscr{M}|$ blocks $B_{(h,j)}$ of dimension $n_h \times m_j$, for $h \in \mathscr{N}$, $j \in \mathscr{M}$.

Remark 2.1 *Since each decision (information) agent can be associated with more columns (rows) of matrix B, in general $n \neq |\mathscr{N}|$ and $m \neq |\mathscr{M}|$. Hence, $m > n$ does not imply that there are more decision agents than information agents. Also, since we are considering partitions, each column (row) of B is associated with a single decision (information) agent.*

Each information agent $h \in \mathscr{N}$ aims at enforcing the constraints $\mathscr{C}_h$: if these constraints are not fulfilled, it imposes an additional penalty, proportional to $B_{(h,:)}u(t) - w_h$, to the cost currently paid by the decision agents. Hence, at each time $t \geq 0$, the decision agents have to pay an overall cost given by the sum of the "basal" cost $\sum_{i=1}^m f_i(u_i)$ and the current penalties imposed by the information agents.

B is a structured matrix, with structural zero-blocks. Agents communicate according to the following rules.

- Each agent can communicate with some agents of the other group, but with no agents of its own group.
- Agent $h \in \mathscr{N}$ and agent $j \in \mathscr{M}$ are *connected* (respectively *not connected*) if the block $B_{(h,j)}$ of matrix B is a non-zero (respectively zero) matrix.
- Each decision agent can communicate only with the information agents connected to it (corresponding to non-zero blocks in its block-column); each information agent can communicate only with the decision agents connected to it (corresponding to non-zero blocks in its block-row) and knows the value of the corresponding components of w.

The connections among decision and information agents can be represented by a bipartite *information-decision graph*, where each edge connecting an information agent $h \in \mathscr{N}$ and a decision agent $j \in \mathscr{M}$ corresponds to a non-zero block $B_{(h,j)}$ of matrix B.

Example 2.1 *Consider the following partitioned matrix B and vectors u and w.*

$$B = \begin{bmatrix} B_{11} & \mathbf{0}_{12} & B_{13} & B_{14} & \mathbf{0}_{15} \\ \mathbf{0}_{21} & \mathbf{0}_{22} & B_{23} & \mathbf{0}_{24} & \mathbf{0}_{25} \\ \mathbf{0}_{31} & \mathbf{0}_{32} & \mathbf{0}_{33} & B_{34} & B_{35} \\ B_{41} & B_{42} & \mathbf{0}_{43} & B_{44} & \mathbf{0}_{45} \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}. \quad (4)$$

We have $|\mathscr{N}| = 4$ and $|\mathscr{M}| = 5$. The corresponding graph is reported in Fig. 1. The structure of B in (4) shows that, for instance, decision agent 1 can communicate only with information agents 1 and 4; information agent 1 can communicate only with decision agents 1, 3 and 4.

We are now ready to state the problem faced in the paper.

Problem 1 *Design an algorithm that allows decision and information agents to solve the optimisation problem* (1)–(2) *in a* decentralised *manner, so that*

- *each decision agent $j \in \mathscr{M}$ sends the current value of the sub-vector u_j only to the information agents connected to it;*
- *each information agent $h \in \mathscr{N}$ sends the current slack values $B_{(h,:)}u - w_h$ of the associated constraints $\mathscr{C}_h$ only to the decision agents connected to it.*

In the next section, the concept of saturation function is introduced and used in the explicit expression of the proposed decentralised mechanism to solve problem (1)–(2).

3 Preliminaries

In view of Assumption 1, the objective function in (1) is strictly convex and continuous in the whole compact and convex domain $\mathscr{D} = \{u \in \mathscr{U} : Bu = w\}$. Hence, an optimal solution u^* always exists and is unique (Boyd & Vandenberghe 2004, pp. 302–304).

Given the bounds (2), the *saturation function* $\text{sat}_{\mathscr{U}}[\cdot] : \mathbb{R}^m \to \mathbb{R}^m$ is component-wise defined as follows (see Fig. 2):

$$\text{sat}_{\mathscr{U}_i}[y_i] \triangleq \begin{cases} u_i^-, & \text{if} \quad y_i < u_i^-, \\ y_i, & \text{if} \quad u_i^- \leq y_i \leq u_i^+, \\ u_i^+, & \text{if} \quad y_i > u_i^+. \end{cases}$$

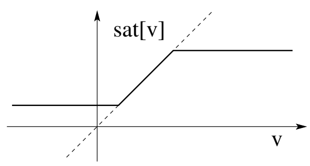

Figure 2. The saturation function.

For $i = 1,\dots,m$, let $g_i \triangleq \frac{df_i}{du_i}$ be the function derivative of f_i. Assumption 1 implies that g_i is strictly increasing, continuous and differentiable. Hence, its inverse function $\phi_i \triangleq$

$g_i^{-1}:\mathbb{R}\to\mathbb{R}$ exists and is increasing. We define the vector function $\phi:\mathbb{R}^m\to\mathbb{R}^m$ mapping $y=[y_1\dots y_m]^\top\in\mathbb{R}^m$ to

$$\phi(y)=[\phi_1(y_1),\phi_2(y_2),\dots,\phi_m(y_m)]^\top. \qquad (5)$$

Lemma 3.1 *There exists a vector $\xi^*\in\mathbb{R}^n$ such that the solution u^* of the optimisation problem (1)–(2) can be expressed as $u^*=sat_{\mathscr{U}}[\phi(-B^\top\xi^*)]$ (namely, $u_i^*=sat_{\mathscr{U}_i}[\phi_i(-B^\top\xi^*)_i]$). Conversely, given a vector $\xi^*\in\mathbb{R}^n$ such that*

$$Bsat_{\mathscr{U}}[\phi(-B^\top\xi^*)]=w, \qquad (6)$$

then $u^*=sat_{\mathscr{U}}[\phi(-B^\top\xi^*)]$ is the solution of (1).

Proof. In the optimisation problem (1), the objective function is a continuously differentiable, convex function over the compact and convex domain $\mathscr{U}$, while the constraints are defined by continuously differentiable affine functions. As a consequence, the Karush-Khun-Tucker (KKT) conditions are necessary and sufficient to identify the unique optimal solution. If ξ, λ and ν denote the vectors of Lagrange multipliers, these conditions are

$$\begin{aligned} g_i(u_i)&=\frac{\partial}{\partial u_i}\left[(w-Bu)^\top\xi-(u^--u)^\top\lambda-(u-u^+)^\top\nu\right]\\ &=-(B^\top\xi)_i+\lambda_i-\nu_i=0,\quad i=1,\dots,m, && (7a)\\ Bu&=w,\quad u^-\le u\le u^+ && (7b)\\ \lambda^\top&(u^--u)=0,\quad \nu^\top(u-u^+)=0 && (7c)\\ \lambda,\nu&\ge 0 && (7d)\\ u,\lambda,\nu&\in\mathbb{R}^m,\quad \xi\in\mathbb{R}^n. && (7e)\end{aligned}$$

Recalling that g_i is strictly increasing, we can rewrite (7a) as

$$u_i=\phi_i(-(B^\top\xi)_i+\lambda_i-\nu_i)\quad i=1,\dots,m \qquad (8)$$

Since an optimal solution u^* exists for problem (1)–(2), then there also exists a solution $(u^*,\xi^*,\lambda^*,\nu^*)$ satisfying the above KKT conditions. The next step is to note that, for $i=1,\dots,m$, this solution satisfies

$$\begin{aligned} u_i^*&=\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)], && (9a)\\ \lambda_i^*&=\max\{g_i(\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)])+(B^\top\xi^*)_i,0\}, && (9b)\\ \nu_i^*&=\max\{-(B^\top\xi^*)_i-g_i(\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)]),0\}. && (9c)\end{aligned}$$

Conditions (7)–(8) trivially imply conditions (9), for all components u_i^* such that $u_i^-<u_i^*<u_i^+$. Indeed, in this case $\lambda_i^*=\nu_i^*=0$ and then $u_i^*=\phi_i(-(B^\top\xi^*)_i+\lambda_i^*-\nu_i^*)=\phi_i(-(B^\top\xi^*)_i)=\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)]$. Now, assume that $u_i^*=u_i^-$ and hence $\lambda_i^*\ge 0$ and $\nu_i^*=0$. Then, (8) implies $u_i^-=\phi_i(-(B^\top\xi^*)_i+\lambda_i^*)$. As $\lambda_i^*\ge 0$ and ϕ_i is increasing, we have $\phi_i(-(B^\top\xi^*)_i)\le u_i^-=\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)]=u_i^*$, that is, condition (9a). Condition (9b) holds as $\lambda_i^*=g_i(u_i^-)+(B^\top\xi^*)_i=g_i(\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)])+(B^\top\xi^*)_i\ge 0$. Finally, condition (9c) holds as $\nu_i^*=0=\max\{-(B^\top\xi^*)_i-g_i(\mathrm{sat}_{\mathscr{U}_i}[\phi_i(-(B^\top\xi^*)_i)]),0\}$. The proof is completed by observing that a symmetric argument holds if $u_i^*=u_i^+$.

Conversely, if ξ^* satisfies the condition (6), then, based on (9), we can build a solution $(u^*,\xi^*,\lambda^*,\nu^*)$ that satisfies the KKT conditions. ■

Lemma 3.2 *Let $\xi^*\in\mathbb{R}^n$ be fixed. There exist nonnegative continuous functions $\Delta_i:\mathbb{R}^n\to\mathbb{R}$, $i=1,\dots,m$, such that, for all $z\in\mathbb{R}^n$,*

$$sat_{\mathscr{U}}[\phi(-B^\top(\xi^*+z))]-sat_{\mathscr{U}}[\phi(-B^\top(\xi^*))]=\Delta(z)(-B^\top z).$$

where $\Delta(z)=diag\{\Delta_i(z)\}$. Moreover, $\|\Delta(z)\|<1/\mu$, for all $z\in\mathbb{R}^m$, where μ is the constant in Assumption 1.

Proof. Let $\psi_i:g_i(\mathbb{R})\to\mathbb{R}$ denote the (non-decreasing) function defined by $\psi_i(\cdot)\triangleq\mathrm{sat}_{\mathscr{U}_i}(\phi_i(\cdot))$, which is differentiable almost everywhere since ϕ is continuous and differentiable, and note that, for any $p,q\in\mathbb{R}$,

$$\psi_i(q+p)-\psi_i(q)=\left[\int_0^1\psi_i'(q+\zeta p)d\zeta\right]p,$$

where ψ' is the right derivative of ψ, since the saturation function is not differentiable. For all i, $\psi_i'=\phi_i'$ when the saturation does not occur, while $\psi_i'=0$ when it occurs. The first claim is then proven by defining

$$\Delta_i(z)=\int_0^1\psi_i'(-(B^\top\xi^*)_i-\zeta(B^\top z)_i)d\zeta.$$

Also, since the derivative of ϕ_i is $\phi_i'=1/g_i'=1/f_i''>0$, in view of Assumption 1, $0\le|\Delta_i|\le|\phi_i'|=|1/f_i''|<1/\mu$. ■

4 Main Results

This section first introduces a continuous-time, dynamical mechanism that provides a decentralised solution to the optimisation problem (1)–(2). Then, a discrete-time scheme is proposed to allow for a numeric implementation.

4.1 Decentralised continuous-time approach

Given the sets of agents $\mathscr{M}$ and $\mathscr{N}$, the decentralised solution to problem (1)–(2) can be obtained as the steady-state value $u^*=\lim_{t\to\infty}u(t)$ generated by the following set of dynamic systems.

Information dynamics: any information agent $h\in\mathscr{N}$ is associated with a dynamic system whose *state variable* $\xi_h\in\mathbb{R}^{n_h}$ evolves according to Eq. (10a) below, starting from an arbitrary initial value $\xi_h(0)=\xi_{h,0}$;

Decision strategy: any decision agent $j\in\mathscr{M}$ is associated with a *control variable* $u_j\in\mathbb{R}^{m_j}$, as in Eq. (10b) below;

Mechanism: the set of dynamic systems evolves according to the equations

$$\dot\xi_h(t)=\sum_{j\in\mathscr{M}}B_{(h,j)}u_j(t)-w_h,\quad\forall h\in\mathscr{N} \qquad (10a)$$

$$u_j(t)=\mathrm{sat}_{\mathscr{U}_j}\left[\phi_j\left(-\sum_{h\in\mathscr{N}}B_{(h,j)}^\top\xi_h(t)\right)\right],\quad\forall j\in\mathscr{M}. \qquad (10b)$$

Remark 4.1 *(Network decentralisation). Conditions (10a)–(10b) allow each of the agents in $\mathcal{N}$ and in $\mathcal{M}$ to locally compute the components of the decentralised solution of their interest. Indeed, each information agent $h \in \mathcal{N}$ determines the value of the associated state variable ξ_h only as a function of the control variables u_j, with $j \in \mathcal{M}$, corresponding to non-zero blocks $B_{(h,j)}$. Similarly, each decision agent $j \in \mathcal{M}$ determines the value of the associated control variable u_j only as a function of the state variables ξ_h, with $h \in \mathcal{N}$, corresponding to non-zero blocks $B^\top_{(h,j)}$.*

Denote by $u \triangleq \left[u_1^\top \; u_2^\top \; \dots \; u_{|\mathcal{M}|}^\top\right]^\top$ and $\xi \triangleq \left[\xi_1^\top \; \xi_2^\top \; \dots \; \xi_{|\mathcal{N}|}^\top\right]^\top$ the vectors associated with the control agents and with the information agents respectively, where u_j are the control variables and ξ_h are the state variables. We now show that u converges to the optimal solution u^* of problem (1)–(2), while $\lim_{t\to\infty} \dot{\xi}(t) = 0$, hence, asymptotically, the information agents will detect no constraint violations. Specifically, we rewrite the system (10) in vector form as

$$\dot{\xi}(t) = Bu(t) - w, \tag{11a}$$
$$u(t) = \mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top \xi(t)\right)\right]. \tag{11b}$$

The following theorem generalises the results by Bauso et al. (2013); Blanchini et al. (2016).

Theorem 4.1 *Under Assumptions 1, 2 and 3, the solution $\xi(t)$ of system (11) is bounded and converges to the set*

$$\Xi = \left\{\xi \in \mathbb{R}^n : B\mathrm{sat}_{\mathcal{U}}\left[\phi(-B^\top \xi)\right] = w\right\},$$

while $u(t)$ converges to the solution u^* of problem (1)–(2).

Proof. Let $\xi^* \in \Xi$ and $z(t) = \xi(t) - \xi^*$. Then,

$$\begin{aligned}\dot{z}(t) &= B\mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top(\xi^* + z(t))\right)\right] - w \\ &= B\left(\mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top(\xi^* + z(t))\right)\right] - \mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top \xi^*\right)\right]\right) \\ &= -B\Delta(z)B^\top z,\end{aligned}$$

where the last step exploits Lemma 3.2. Consider the Lyapunov function $V(z) = \frac{1}{2} z^\top z$, whose Lyapunov derivative is

$$\dot{V}(z) = -z^\top B\Delta(z)B^\top z \leq 0,$$

because $\Delta(z)$ is a diagonal matrix of nonnegative functions. So, z is bounded and, in view of LaSalle's principle, converges to the set where $\dot{V}(z) = -z^\top B\Delta(z)B^\top z = 0$. Now, for any symmetric positive (or negative) semidefinite matrix S, $z^\top S z = 0$ if and only if $Sz = 0$. Hence, z converges to the set

$$\begin{aligned}\mathcal{Z} &= \{z \in \mathbb{R}^n : B\Delta(z)B^\top z = 0\} \\ &= \left\{z \in \mathbb{R}^n : \; B\mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top(\xi^* + z)\right)\right] - w = 0\right\},\end{aligned} \tag{12}$$

hence $\xi(t)$ converges to the set Ξ. In view of the continuity and of Lemma 3.1, $u(t)$ converges to the optimum u^*. ■

Remark 4.2 *While u^* is unique, the value ξ^* such that $B\mathrm{sat}_{\mathcal{U}}\left[\phi(-B^\top \xi^*)\right] = w$ may be not unique (hence, the set $\mathcal{Z}$ may include also non-zero vectors). As an example of non-uniqueness, if $B = [1\; 1]$, $1 \leq u_1 \leq 2$, $3 \leq u_2 \leq 4$, $\phi(x) = x$ and $w = 5$, then any $\xi^* \in [-3, -2]$ is suitable. However, if $\mathcal{Z} = \{0\}$ is a singleton, then necessarily $z(t) \to 0$.*

Theorem 4.1 helps us prove the convergence of the discrete-time algorithm described below.

4.2 Decentralised discrete-time algorithm

System (11) can be implemented through a numerical algorithm where the state equation (11a) is discretised according to an Euler scheme, with sampling time $\tau > 0$, to obtain:

$$\xi(k+1) = \xi(k) + \tau Bu(k) - \tau w, \tag{13a}$$
$$u(k) = \mathrm{sat}_{\mathcal{U}}\left[\phi\left(-B^\top \xi(k)\right)\right]. \tag{13b}$$

Clearly, this discrete-time algorithm preserves the decentralised nature of the continuous-time solution.

We can show that, for τ sufficiently small, the discrete-time sequence $\xi(k)$ defined by (13) converges to ξ^*; hence, the control sequence $u(k)$ converges to u^*.

Theorem 4.2 *Under Assumptions 1, 2 and 3, and if*

$$\tau < \frac{2\mu}{\|B\|^2}, \tag{14}$$

where μ is the constant in Assumption 1, then the sequence $u(k)$, which evolves according to system (13), converges to u^*, solution of problem (1)–(2).

The proof of Theorem 4.2 requires two lemmas.

Lemma 4.1 *Given a symmetric, positive semidefinite matrix S, if $0 < \tau < 2/\|S\|$, then $\|I - \tau S\| \leq 1$.*

Proof. Being $S \succeq 0$, its eigenvalues λ_k are in the interval $[0, \|S\|]$. The eigenvalues of $I - \tau S$ are $1 - \tau\lambda_k$ and, if $0 < \tau < 2/\|S\|$, they lie in the interval $(-1, 1]$. Then, since $I - \tau S$ is symmetric as well, its norm must be less or equal to 1. ■

Lemma 4.2 *Given a symmetric positive semidefinite matrix $S \in \mathbb{R}^{m\times m}$ and a vector $z \in \mathbb{R}^m$, if $0 < \tau < 2/\|S\|$, then $\|z\| = \|(I - \tau S)z\|$ if and only if $Sz = 0$.*

Proof. Clearly $Sz = 0$ implies the norm equality. To prove the opposite, take an orthonormal matrix Q such that $Q^\top S Q = \Sigma = \mathrm{diag}\{\mathbf{0}, \Sigma_2\}$, where Σ_2 is a diagonal matrix with the positive eigenvalues of S and $\mathbf{0}$ is the null matrix associated with the zero eigenvalues. Let $\hat{z} = Q^\top z$. Being Q orthonormal, it does not change the norm: $\|\hat{z}\| = \|z\|$. Then, $\|z\| = \|(I - \tau S)z\|$ is equivalent to $\|\hat{z}\| = \|(I - \tau Q^\top S Q)\hat{z}\|$. Partitioning the components and squaring gives

$$\left\|\begin{bmatrix}\hat{z}_1 \\ \hat{z}_2\end{bmatrix}\right\|^2 = \left\|\begin{bmatrix}\hat{z}_1 \\ \hat{z}_2\end{bmatrix} - \begin{bmatrix}0 & 0 \\ 0 & \tau\Sigma_2\end{bmatrix}\begin{bmatrix}\hat{z}_1 \\ \hat{z}_2\end{bmatrix}\right\|^2.$$

In view of the property of the Euclidean norm we have

$$\|\hat{z}_1\|^2 + \|\hat{z}_2\|^2 = \|\hat{z}_1\|^2 + \|(I - \tau\Sigma_2)\hat{z}_2\|^2,$$

namely, $\|\hat{z}_2\|^2 = \|(I - \tau\Sigma_2)\hat{z}_2\|^2$. Since Σ_2 includes the positive eigenvalues only, the eigenvalues of the diagonal matrix $(I - \tau\Sigma_2)$ lie in the open interval $(-1, 1)$ if $0 < \tau < 2/\|S\|$, hence their magnitude is strictly less than 1. Therefore, the norm equality is possible only if $\hat{z}_2 = 0$. Then $\Sigma\hat{z} = 0$ and $Sz = Q\Sigma Q^\top z = Q\Sigma\hat{z} = 0$. Hence, $Sz = 0$. ■

Proof of Theorem 4.2. Let $z(k) = \xi(k) - \xi^*$ and, as done for the continuous-time case, write the system as

$$z(k+1) = \left(I - \tau B\Delta(z(k))B^\top\right) z(k). \tag{15}$$

In view of (14) and Lemma 3.2 we have

$$\tau < \frac{2\mu}{\|B\|^2} \leq \frac{2}{\|B\Delta(z(k))B^\top\|}.$$

Then, since $\|I - \tau B\Delta(z(k))B^\top\| \leq 1$ in view of Lemma 4.1,

$$\|z(k+1)\| \leq \|I - \tau B\Delta(z(k))B^\top\|\|z(k)\| \leq \|z(k)\|.$$

Therefore, $z(k)$ is bounded.

To prove convergence, we invoke LaSalle's invariance principle for nonlinear discrete-time dynamical systems (Sundarapandian 2003). Since the function $\|z(k+1)\|$ is non-increasing, necessarily $z(k)$ converges to the set for which

$$\|z\| = \|(I - \tau B\Delta(z)B^\top)z\|.$$

According to Lemma 4.2, this is exactly the set for which $B\Delta(z)B^\top z = 0$, namely, the set $\mathscr{Z}$ in (12). As in Theorem 4.1, we can therefore conclude that $u(k)$ converges to u^*. ■

Next, we show that condition (14) is crucial, since it becomes also necessary for convergence if the constraints are not active and the function to be minimised is proportional to the Euclidean norm of u.

Proposition 1 *Take $f_i(u_i) = \frac{1}{2}u_i^2$ and assume $u^* \in int\mathscr{U}$. Then, the discrete-time system* (13) *converges only if* (14) *holds.*

Proof. In a neighborhood of u^*, the saturation is not active, $\Delta = I$, and equation (15) becomes $z(k+1) = (I - \tau BB^\top)z(k)$. The eigenvalues of matrix $(I - \tau BB^\top)$ are $1 - \tau\lambda_i$, where λ_i are the (positive) eigenvalues of $BB^\top$. The eigenvalues of matrix $B^\top B$ are those of matrix $BB^\top$ plus $m - n$ zero eigenvalues; in both cases, the maximum eigenvalue is equal to $\|B\|^2$. Then, condition (14) is clearly necessary for $|1 - \tau\lambda_i| < 1$ which, in turn, is necessary for convergence. Note that here $\mu = 1$. ■

Remark 4.3 *When $f_i(u_i) = \frac{1}{2}u_i^2$, the maximum value of the time-discretisation τ is proportional to the inverse of the largest eigenvalue of matrix $BB^\top$ (or, equivalently, the largest eigenvalue of matrix $B^\top B$, being both equal to $\|B\|^2$). The value of τ can be optimally chosen by considering the non-saturated case, when $u = -B^\top\xi$ and convergence depends on the eigenvalue of matrix $(I - \tau BB^\top)$. Convergence can be optimised by minimising the eigenvalue with maximum modulus: $\min\limits_{\tau>0} \max\{|1 - \tau\lambda_M|, |1 - \tau\lambda_m|\}$, where λ_m and λ_M are, respectively, the smallest and the largest eigenvalue of $BB^\top$. The optimal value turns out to be the inverse of the average between the minimum and the maximum eigenvalue of $BB^\top$:*

$$\tau^* = \frac{2}{\lambda_M + \lambda_m}.$$

4.3 Speed of convergence of the algorithm

Here we show that, under mild conditions, the convergence of the algorithm is exponential. To simplify the exposition, we consider $f_i(u_i) = \frac{1}{2}u_i^2$, so that $u = \text{sat}_{\mathscr{U}}\left[-B^\top\xi\right]$.

Theorem 4.3 *Assume that (i) there is a unique vector ξ^* such that $u^* = sat_{\mathscr{U}}\left(-B^\top\xi^*\right)$ and (ii) at least $r \geq n$ components of u^* are not saturated (namely, $u_i^- < u_i^* < u_i^+$) and the corresponding r columns of B span $\mathbb{R}^n$. Then, for $z = \xi - \xi^*$ and $V(z) = \|z\|^2$, there exists $\beta > 0$ such that the solution of system* (11) *satisfies*

$$\dot{V}(z(t)) \leq -2\beta^2 V(z(t)), \tag{16}$$

for all $t \geq \theta(\xi(0)) > 0$, where $\theta(\xi(0))$ is a time value depending on the initial condition $\xi(0)$.

Proof. System (11) can be equivalently written as

$$\begin{aligned}\dot{z}(t) &= B\left(\text{sat}_{\mathscr{U}}\left[-B^\top(z + \xi^*)\right] - \text{sat}_{\mathscr{U}}\left[-B^\top\xi^*\right]\right)\\ &\triangleq B\sigma\left[-B^\top z\right],\end{aligned}$$

where σ is a new saturation function with translated bounds $u_i^- - \text{sat}_{\mathscr{U}_i}\left[-(B^\top\xi^*)_i\right]$ and $u_i^+ - \text{sat}_{\mathscr{U}_i}\left[-(B^\top\xi^*)_i\right]$ instead of u_i^- and u_i^+. Note that zero is inside these new bounds.

By assumption, there are at least r non-saturated components at steady state; assume they are $u_{ns} = [u_1^* \dots u_r^*]^\top$, where ns stands for "non-saturated". For these components, there is a neighborhood $\mathscr{Z}_0$ of $z = 0$ where $\sigma_i[-(B^\top z)_i] = (-B^\top z)_i$. Denote by $u_s = [u_{r+1}^* \dots u_m^*]^\top$ the other (possibly saturated) components. Since $\xi(t) \to \xi^*$, and hence $z(t) \to 0$, for any $\xi(0)$, there exists $\theta(\xi(0)) > 0$ such that $z(t) \in \mathscr{Z}_0$ for $t > \theta(\xi(0))$. As a consequence, the sub-vector $u_{ns}(t)$ is not saturated. For the other components,

$$\sigma_i[-(B^\top z)_i] = \Delta_i(z)(-B^\top z)_i, \quad 0 \leq \Delta_i(z) \leq 1,$$

for some function $\Delta_i(z)$. By grouping the $\Delta_i(z)$'s in a diagonal matrix $\Delta^s(z)$, the overall system for $t > \theta(\xi(0))$ becomes

$$\dot{z}(t) = -[B_{ns}B_{ns}^\top + B_s\Delta^s(z(t))B_s^\top]z(t), \tag{17}$$

where B_{ns} and B_s consist of the columns of B associated with non-saturated and saturated components, respectively. Since the columns of B_{ns} span $\mathbb{R}^n$, $B_{ns}B_{ns}^\top$ is positive definite. The derivative of $V(z) = z^\top z$ is

$$\dot{V}(z) = -2z^\top B_{ns}B_{ns}^\top z - 2z^\top B_s \Delta^s(z) B_s^\top z \leq -2z^\top B_{ns}B_{ns}^\top z \\ \leq -2\beta^2 z^\top z = -2\beta^2 V(z),$$

where β is the smallest singular value of matrix B_{ns}. ■

Remark 4.4 *The condition of Theorem 4.3 implies that the difference* $\xi - \xi^*$ *converges to zero as fast as* $e^{-\beta t}$. *Assumptions (i)-(ii), which are not demanding and are generically satisfied, ensure* exponential *convergence, while convergence is always guaranteed as long as Assumptions 1–3 are satisfied. Exponential convergence is not achievable, in general, in the presence of saturations (Hu & Lin 2001).*

The following corollary addresses the performance of the discrete-time algorithm.

Corollary 4.1 *Under the same hypotheses of Theorem 4.3, the discrete-time algorithm converges exponentially if* $0 < \tau < 2/\|B\|^2$.

Proof. As done in the proof of Theorem 4.3, we absorb the system in the linear differential inclusion (17), where $\Delta^s(z(t))$ is a diagonal matrix whose nonnegative diagonal entries are bounded by 1. As shown in the theorem proof, the linear differential inclusion (17) is stable. Then, consider the corresponding Euler system

$$z(k+1) = \left[I - \tau(B_{rs}B_{rs} + B_s\Delta^s(z(k))B_s^\top)\right] z(k) .$$

Since the diagonal matrix Δ^s has positive entries upper bounded by 1, $\|[B_{ns}B_{ns}^\top + B_s\Delta^s B_s^\top]\| \leq \|BB^\top\|$. Then, if $0 < \tau < 2/\|B\|^2$, convergence is ensured. On the other hand, if a differential inclusion converges, then it converges exponentially (Blanchini & Miani 2015). ■

4.4 Dealing with uncertain models

As a step toward robust optimisation (Ben-Tal, El Ghaoui & Nemirovski 2006; Bertsimas & Thiele 2006), we consider the case in which B is characterised by an additive uncertainty ΔB. In this case, while the decision agents (u) base their strategy on the nominal B, the information agents (ξ) measure and integrate the true constraint violation $(B + \Delta B)u - w$. As a consequence, equation (11b) remains unchanged, while equation (11a) must be replaced by

$$\dot{\xi}(t) = (B + \Delta B)u(t) - w . \tag{18}$$

When $f_i(u_i) = \frac{1}{2}u_i^2$, we can show that convergence properties are preserved also when B is uncertain, provided that a standard assumption in robust control is satisfied.

Assumption 4 *For all* $\ell = 1, \ldots, m$, *denoting by* $\underline{B^\top}_\ell$ *the* ℓ*-th row of* $B^\top$ *and by* $(B + \Delta B)|_\ell$ *the* ℓ*-th column of* $B + \Delta B$,

$$\underline{B^\top}_\ell (B + \Delta B)|_\ell = \sum_{i=1}^{n} B^\top_{(\ell,i)} (B + \Delta B)_{(i,\ell)} \geq 0 .$$

The assumption means that the inner product of the nominal column $B|_\ell$ and the actual column $(B + \Delta B)|_\ell$ cannot be negative, otherwise the effect of the input u_ℓ would be the opposite of the intended one. Typically, $B|_\ell$ has only a subset of structurally non-zero entries and Assumption 4 is satisfied as long as their sign is preserved in $(B + \Delta B)|_\ell$, which is a reasonable assumption on the magnitude of the uncertainty.

Theorem 4.4 *Consider the system defined by (11b) and (18). Under Assumptions 2, 3 and 4, the trajectory* $\xi(t)$ *is bounded. Moreover, if there exists only one constant vector* ξ^* *such that* $(B + \Delta B)sat_{\mathscr{U}}[-B^\top \xi^*] = w$, *then* $\xi(t)$ *converges to* ξ^* *and* $u(t)$ *converges to the solution* u^* *of the optimisation problem*

$$\min_{u \in \mathscr{U}} \quad \frac{1}{2}u^\top u \\ \text{s.t.} \quad (B + \Delta B)u = w ,$$

with the feasible set $\mathscr{U}$ *defined as in (2).* □

Proof. By taking $z(t) = \xi(t) - \xi^*$, let us consider the candidate Lyapunov function $V(z) = \frac{1}{2}z^\top z$, which is positive definite and radially unbounded. Its time-derivative along the system trajectories is

$$\begin{aligned} \dot{V}(z) &= z^\top[(B + \Delta B)u(t) - w] \\ &= z^\top(B + \Delta B)(\text{sat}_{\mathscr{U}}[-B^\top(z + \xi^*)] - \text{sat}_{\mathscr{U}}[-B^\top \xi^*]) \\ &= \sum_{\ell=1}^{m} z^\top(B + \Delta B)|_\ell(\text{sat}_{\mathscr{U}}[-\underline{B^\top}_\ell(z + \xi^*)] - \text{sat}_{\mathscr{U}}[-\underline{B^\top}_\ell \xi^*]) . \end{aligned}$$

In view of Assumption 4, for $\ell = 1, \ldots, m$, the inner product of $\underline{B^\top}_\ell$ and $(B + \Delta B)|_\ell$ is nonnegative, and is the only (possibly) non-zero eigenvalue of the $\mathbb{R}^{n \times n}$ diagonalisable matrix $B_\ell^\Delta \triangleq (B + \Delta B)|_\ell \underline{B^\top}_\ell$. So, for all $y \in \mathbb{R}^n$, we have $y^\top B_\ell^\Delta y \geq 0$. Taking $y = z$, we obtain

$$z^\top(B + \Delta B)|_\ell \underline{B^\top}_\ell z \geq 0 , \tag{19}$$

which implies that either one of the two scalars $z^\top(B + \Delta B)|_\ell$ and $\underline{B^\top}_\ell z$ is zero or they have the same sign. In both cases,

$$z^\top(B + \Delta B)|_\ell(\text{sat}_{\mathscr{U}}[-\underline{B}_\ell^\top(z + \xi^*)] - \text{sat}_{\mathscr{U}}[-\underline{B}_\ell^\top \xi^*]) \leq 0,$$

because, for two scalars x and y, $(y - x)(\text{sat}_{\mathscr{U}}[x] - \text{sat}_{\mathscr{U}}[y]) \leq 0$ (take $x = -\underline{B^\top}_\ell \xi$ and $y = -\underline{B^\top}_\ell \xi^*$). The proof can be concluded by resorting to LaSalle's invariance principle. ■

Remark 4.5 *With the same considerations made in the previous subsections, it is possible to show that the discrete-time algorithm converges even in the presence of uncertainties, provided that* $\tau > 0$ *is small enough.*

5 Dynamic Network-Decentralised Optimal Control

The proposed dynamic mechanism can be exploited to design optimal network-decentralised control strategies (Ataslar & Iftar 1998; Iftar 1999; Iftar & Davison 2002;

Blanchini, Franco & Giordano 2013, 2015). Consider a discrete-time system of the form

$$x(k+1) = Fx(k) + G\omega(k)\,, \tag{20}$$

where $F \in \mathbb{R}^{n\times n}$ is block-diagonal, $F = \mathrm{diag}\{F_1, F_2, \ldots, F_N\}$, while $G \in \mathbb{R}^{n\times m}$ is a suitably block-structured matrix with N block-rows and M block-columns. Each block $G_{(i,j)}$, $i = 1,\ldots,N$, $j = 1,\ldots,M$, has the same number of rows as F_i. Several blocks $G_{(i,j)}$ are structurally zero, hence we seek for a network-decentralised control strategy for system (20) that exploits the sparsity structure of G, according to the framework by Blanchini, Franco & Giordano (2013, 2015). In particular, we look for the sequence of control inputs in the interval $[0,T]$ that, given the initial state $x(0)$, minimises a positive combination of the inputs norm and of the final state norm, with inputs subject to interval constraints:

$$\begin{aligned}
\min_{x(T),\omega} \quad & \eta\|x(T)\|^2 + \sum_{k=0}^{T-1}\|\omega(k)\|^2 \\
\text{s.t.} \quad & x(T) = F^T x(0) + \sum_{k=0}^{T-1} F^{T-k-1}G\omega(k), \\
& \omega^- \le \omega(k) \le \omega^+, \quad k = 0,\ldots,T-1
\end{aligned} \tag{21}$$

with $\eta > 0$. If we denote the decision variable vector as

$$u = [x(T)^\top\ \omega(T-1)^\top\ \ldots\ \omega(1)^\top\ \omega(0)^\top]^\top,$$

the problem can be cast in the form (1)–(2), with functions $f_i(u_i) = \eta u_i^2$ for $i = 1,\ldots,n$ and $f_i(u_i) = u_i^2$ for $i > n$,

$$\tilde{B} = [-I\ \ G\ \ FG\ \ F^2G\ \ldots\ F^{T-1}G] \quad \text{and} \quad w = -F^T x(0)\,.$$

Denote by $G_{(:,i)}$ the ith block-column of G and define $E_i = [G_{(:,i)}\ \ FG_{(:,i)}\ \ F^2G_{(:,i)}\ \ldots\ F^{T-1}G_{(:,i)}]$. We can obtain a new constraint matrix B by rearranging the columns of $\tilde{B}$ as follows:

$$B = [-I\ \ E] = [-I\ \mid\ E_1\ \ E_2\ \ \ldots\ \ E_M], \tag{22}$$

where now $B = [-I\ \ E]$ *has the same sparsity structure as* G, because F^k is block-diagonal (hence the zero block-rows of $G_{(:,i)}$ correspond to zero block-rows of E_i), and the identity does not add any coupling.

Then, the optimal control sequence can be decided in a decentralised way according to the following theorem.

Theorem 5.1 *Given a block-diagonal matrix F and a block structured matrix G, the optimal control problem* (21) *can be solved by a network-decentralised optimisation scheme $\dot{\xi}(t) = Bu(t) - w$, with $u(t) = sat[-\phi(B^\top\xi(t))]$ and B as in* (22)*, where the decision agents iteratively compute u, the information agents compute the integral variables ξ associated with the constraints $Bu = w$ and the decision-information graph structure is given by G.*

Remark 5.1 *The "state variable" of the solution algorithm, $\xi(t)$, introduced to guarantee constraint satisfaction, and the state variable of the plant to be controlled, $x(k)$, should not be confused. When ξ reaches the steady state, as it is guaranteed to happen, the corresponding value of u is the optimal control sequence on the chosen horizon, which can be applied (open-loop) to the plant. The algorithm must work on a faster time-scale than the plant. To achieve a feedback scheme, in a model predictive fashion, we can compute the control sequence, apply only the first input, measure the new state and recompute the sequence, at every time instant k.*

The cost in (21) only penalises the final state and the control action. However, the scheme can be adapted to a more general setup in which all the states are penalised, just by considering the new cost function

$$\sum_{h=1}^{T}\eta\|x(h)\|^2 + \sum_{k=0}^{T-1}\|\omega(k)\|^2$$

and the extended set of equality constraints related to (20):

$$\left[\begin{array}{ccccc|cccc}
-I & 0 & \cdots & 0 & 0 & G & 0 & \cdots & 0 \\
F & -I & \cdots & 0 & 0 & 0 & G & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & F & -I & 0 & 0 & \cdots & G
\end{array}\right]
\begin{bmatrix} x \\ \omega \end{bmatrix} =
\begin{bmatrix} -Fx(0) \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where $[x^\top\ \ \omega^\top]^\top = [x(1)^\top \ldots x(T)^\top \omega(0)^\top \ldots \omega(T-1)^\top]^\top$. If F and G are block-structured, by rearranging the blocks, we get a block-structured optimisation problem, which can be solved using the proposed decentralised algorithm.

6 Applications

6.1 Strategic blending problem

In this section a *strategic blending* problem is considered. In particular, a deterministic and static version of the problem is analysed in detail; however, a similar analysis can be performed analogously in the dynamic and uncertain case.

The problem can be described as the optimal supply of a good (Sarimveisa et al. 2008; Silver & Peterson 1985), e.g. fuel, from a number of stocking places to a number of destinations, guaranteeing that at each destination the good exhibits a target level of each of a number of characteristics. These characteristics can be chemical-physical, such as, in the case of fuel, the purity or the number of octanes, or geographical, such as the price of transportation. The complexity of the problem is increased by the fact that, with respect to each of the characteristics, the good, or "raw material", of two different stocking places may have two different levels; as a consequence, a blending process has to be designed at each destination to comply with the desired level. Let $\mathscr{S}$, $\mathscr{R}$ and $\mathscr{K}$ denote the set of destinations, of stocking places and of characteristics, respectively. The parameters (i.e., given data) of the problem are:

b_{krs}: amount of characteristic k, per unit of mass of raw material (coming from the stocking place) r, reaching s;
w_r: availability of raw material (at the stocking place) r;
$\hat{w}_s$: demand of good at destination s;

$\tilde{w}_{ks}$: desired level of characteristic k that should be satisfied by good at destination s;
$\tilde{u}^+_{ks}$, $\tilde{u}^-_{ks}$: bounds for allowed deviation from the desired level $\tilde{w}_{ks}$.

The control variables are:

u_{rs} : amount of raw material r that reaches destination s;
$\hat{u}_r$: unused stock of raw material r;
$\tilde{u}_{ks}$: deviation from the desired level $\tilde{w}_{ks}$.

The constraints are:

- The available amount of each raw material is either used or left in stock:
$$\sum_{s\in\mathscr{S}} u_{rs} + \hat{u}_r = w_r\,,\ \forall r \in \mathscr{R}\,. \tag{23}$$
- No more raw material than the available amount can be used:
$$0 \le \hat{u}_r \le w_r\,, \qquad \forall r \in \mathscr{R} \text{ and } \forall s \in \mathscr{S}\,, \tag{24}$$
$$0 \le u_{rs} \le \min(w_r, \hat{w}_s)\,, \quad \forall r \in \mathscr{R} \text{ and } \forall s \in \mathscr{S}\,. \tag{25}$$
- At each destination the demand is filled:
$$\sum_{r\in\mathscr{R}} u_{rs} = \hat{w}_s \forall s \in \mathscr{S}\,. \tag{26}$$
- At each destination the level of each characteristic k must be satisfied within a given level of tolerance:
$$\sum_{r\in\mathscr{R}} b_{krs}u_{rs} + \tilde{u}_{ks} = \tilde{w}_{ks}\,, \quad \forall k \in \mathscr{K} \text{ and } \forall s \in \mathscr{S}\,, \tag{27}$$
$$\tilde{u}^-_{ks} \le \tilde{u}_{ks} \le \tilde{u}^+_{ks}\,, \qquad \forall k \in \mathscr{K} \text{ and } \forall s \in \mathscr{S}\,. \tag{28}$$

As far as the objective function is concerned, the amount of raw materials that is moved, the unused stocks and the deviations from the desired levels of each characteristic should be minimised. As a consequence, the cost function is
$$C(u_{rs}, \hat{u}_r, \tilde{u}_{ks}) \triangleq \sum_{s\in\mathscr{S}}\sum_{r\in\mathscr{R}} c_{rs}u^2_{rs} + \sum_{s\in\mathscr{S}}\sum_{k\in\mathscr{K}} \tilde{c}_{ks}\tilde{u}^2_{ks} + \sum_{r\in\mathscr{R}} \hat{c}_r\hat{u}^2_r \tag{29}$$
where:

c_{rs}: cost per squared unit mass for the transportation of raw material r to the destination s;
$\tilde{c}_{ks}$: cost per squared unit mass for the deviation from the desired level of characteristic k at the destination s;
$\hat{c}_r$: cost for unused raw material r per squared unit mass;

We can cast the problem within our theoretical framework and find the solution as the equilibrium state of a dynamical system where information and decision agents are as follows.

- **Information agents**: one agent is associated with each raw material (constraints (23)) and with each destination (constraints (26) and (27)); hence, $\mathscr{N} = \mathscr{R} \cup \mathscr{S}$;
- **Decision agents**: one agent is associated with each element of $\mathscr{R}\times\mathscr{S}$ for managing each possible transportation of raw materials to a destination (variables u_{rs}); one agent is introduced for managing each unused raw materials (variables $\hat{u}_r$); one agent is associated with each element of $\mathscr{S}\times\mathscr{K}$ for managing each deviations at each destination (variables $\tilde{u}_{kr}$); hence, $\mathscr{M} = (\mathscr{R}\times\mathscr{S}) \cup \mathscr{R} \cup (\mathscr{S}\times\mathscr{K})$.

Table 1
Desired level of characteristics at destinations. The columns report: Destination; Demand; Octane Rating (p.u.); Octane Rating.

$s \in \mathscr{S}$	$\hat{w}_s$	$\tilde{w}_{\alpha s/u}$	$\tilde{w}_{\alpha s}$
A	7000	85	595000
B	6000	93	558000

Table 2
Availability and characteristics transportation. The columns report: Stock; Availability; Octane Rating; Price for $s = A$; Price for $s = B$.

$r \in \mathscr{R}$	w_r	$b_{\alpha rA} = b_{\alpha rB}$	$b_{\beta rA}$	$b_{\beta rB}$
1	2000	70	9.0	2
2	4000	80	12.5	8
3	4000	85	12.5	8
4	5000	90	27.5	12
5	3000	99	27.5	15

Table 3
Allowed deviations from the desired levels. The columns report: Destination; Lower bound Octane Rating; Upper bound Octane Rating; Lower Bound Price; Upper Bound Price.

$s \in \mathscr{S}$	$\tilde{u}^-_{\alpha s}$	$\tilde{u}^+_{\alpha s}$	$\tilde{u}^-_{\beta s}$	$\tilde{u}^+_{\beta s}$
A	-29750	29750	-338000	338000
B	-27900	27900	-179000	179000

By normalising the cost, which is equivalent to suitably rescaling the control variables, (29) can be transformed into $C(u) = \frac{1}{2}u^\top u$, where u is the vector of all decision agents; namely, adopting the notation in (1), $f_i(u_i) = \frac{1}{2}u_i^2$. This leads to the optimisation problem
$$\begin{aligned} \min_{u\in\mathscr{U}} \quad & \frac{1}{2}u^\top u \\ \text{s.t.} \quad & Bu - w = 0\,, \end{aligned}$$
where $Bu = w$ is achieved by grouping the linear constraints (23), (26) and (27), thus forming the sparse matrix B consistently with the information-decision structure previously described and the data vector w, while (24), (25) and (28) correspond to the constraints $u^- \le u \le u^+$, having the form (2). Introducing the integral variable ξ, with $\dot{\xi} = Bu - w$, ensures that the decentralised algorithm $u = \text{sat}[-B^\top\xi]$ drives $\dot{\xi}(t)$ to zero asymptotically, leading to the optimal u^*.

We have numerically simulated a particular instance of this problem, concerning the optimal distribution of fuel with two characteristics from five sources to two destinations: $\mathscr{S} = \{A, B\}$ (destinations), $\mathscr{R} = \{1, 2, 3, 4, 5\}$ (stocks) and $\mathscr{K} = \{\text{Octane Rating}, \text{Price}\} = \{\alpha, \beta\}$ (characteristics). The parameters specifying the problem are reported in Tables 1-3, while all the costs c_{rs}, $\tilde{c}_{ks}$ and $\hat{c}_r$ are set to 1.

Moreover we assume:
$\tilde{u}_{\alpha s} \in [-0.05\cdot\tilde{w}_{\alpha s}, 0.05\cdot\tilde{w}_{\alpha s}]$, a maximum $\pm 5\%$ variation with respect to α;
$\tilde{u}_{\beta s} \in \left[-\min(w_r, \hat{w}_s)\cdot b_{\beta rs}, \min(w_r, \hat{w}_s)\cdot b_{\beta rs}\right]$, an upper

bound on u_{rs} due to transportation cost: $\tilde{w}_{\beta s} = [0, 0]^\top$.

In this way, matrix B and vector w are determined according to the values in Tables 1-3 and the previous assumptions. Hence, one can compute $\lambda_M \simeq 3.8 \times 10^4$, $\lambda_m \simeq 3.5 \times 10^{-3}$ and, consequently, $\bar{\tau} \simeq 5.2 \times 10^{-5}$. The simulation results are reported in Figures 3–4. Fig. 3 (top) reports the discrete-time evolution of the Lyapunov function, with a decreasing behaviour, Fig. 3 (bottom) shows the variation of the input variables through the iterations, while Fig. 4 shows, for each stocking place, the amount of raw material sent to each of the two destinations or left in the stock.

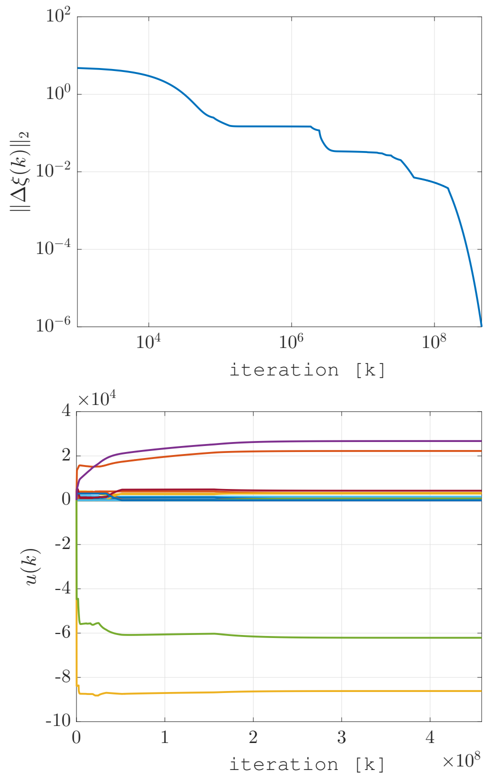

Figure 3. Discrete-time evolution for the strategic blending problem in Section 6.1. Top: Lyapunov function. Bottom: inputs $u(k)$.

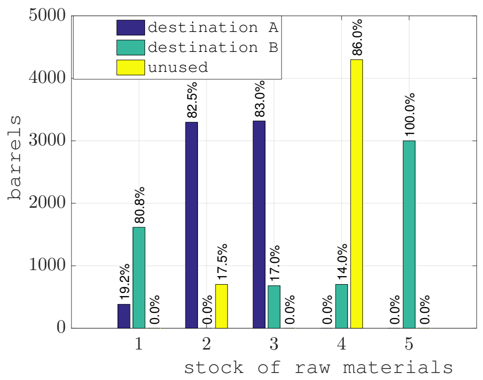

Figure 4. Blending distribution for the problem in Section 6.1.

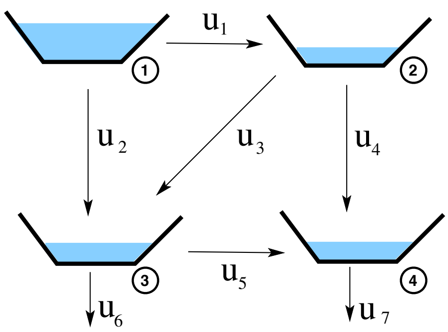

Figure 5. Sketch of the flood control problem in Section 6.2.

6.2 *Flood control problem.*

The devised algorithm for decentralised optimisation is here applied to compute a *network-decentralised* optimal controller that brings a system of reservoirs back to an equilibrium after a flood. For this problem, none of the available model predictive strategies (Breckpot, Agudelo & De Moor 2013; Breckpot et al. 2013; Delgoda et al. 2013; Montero et al. 2013) follows a decentralised approach as the one we propose herein. Consider the network of reservoirs in Fig. 5, where, besides the exchange of fluid between reservoirs, each reservoir has a natural outflow. This scenario can be modelled as in equation (20), with $F = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ and

$$G = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & -1 \end{bmatrix}.$$

We can compute a network-decentralised optimal control strategy on the horizon $[0, T]$ by solving problem (21), where F and G are the matrices reported above, with $\lambda_1 = 0.9$, $\lambda_2 = 1$, $\lambda_3 = 0.95$, $\lambda_4 = 0.9$, while $\omega_i^- = 0$ and $\omega_i^+ = 0.1$ for all i, and $\eta = 1$. We take the initial condition $x(0) = [10 \quad 2 \quad 3 \quad 4]^\top$, which models a flood in reservoir 1, and simulate the evolution of the proposed scheme, with a receding horizon of $T = 12$ steps. The achieved final state is $x(T) = [1.6477 \ 1.2727 \ 1.3038 \ 1.2833]^\top$, while the time evolutions of the buffer levels and the input sequence are reported in Fig. 6. With $\tau = 0.0228$, the procedure converges in about 400 steps and the computation requires 0.003 seconds on a standard PC (clock frequency 2.3 GHz). The optimal control sequence is reported in Table 4. The components u_1 and u_2 are initially small because, due to the high level of the first tank, there is a strong natural outflow $\lambda_1 x_1$; they saturate later, when the natural outflow becomes smaller. Yet, if $\lambda_1 = 1$ (i.e., no natural outflow), then u_1 and u_2 saturate from the very beginning.

In the considered problem, the final state $x(T)$ is a decision variable, whose value is determined by solving the optimisation problem. However, it can also be imposed as a target: this leads to a different problem, where $x(T) = \bar{x}(T)$ is a constraint. For instance, the desired final state $\bar{x}(T) = [1 \ 1 \ 1 \ 1]^\top$ can be imposed over the horizon $T = 12$, resulting in a dif-

Table 4
Optimal control sequence for the flood control problem in Section 6.2.

t	1	2	3	4	5	6	7	8	9	10	11	12
u_1	0	0	0	0.0061	0.0371	0.0715	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
u_2	0.0305	0.0440	0.0595	0.0773	0.0977	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
u_3	0.0999	0.0908	0.0812	0.0712	0.0606	0.0494	0.0376	0.0253	0.0122	0	0	0
u_4	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.0868	0.0662	0.0432	0.0177	0
u_5	0.0839	0.0831	0.0817	0.0796	0.0766	0.0727	0.0678	0.0616	0.0539	0.0447	0.0336	0.0205
u_6	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
u_7	0.0889	0.0988	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000

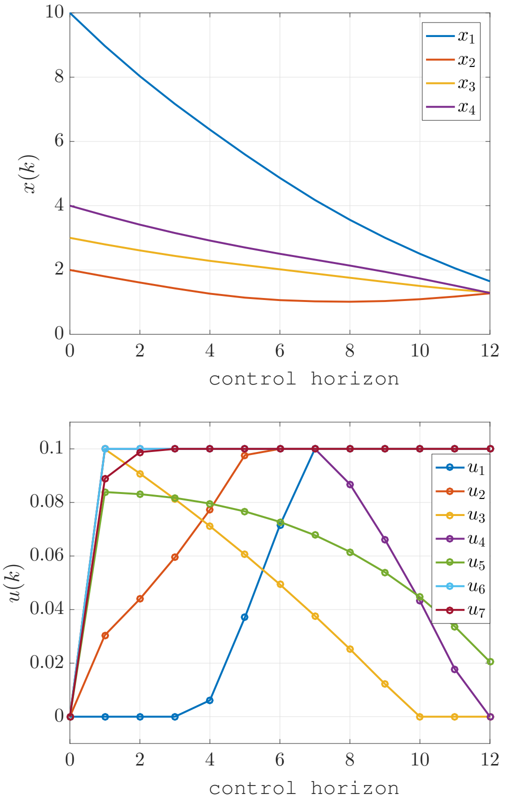

Figure 6. The transient behaviour in terms of buffer levels x_i (top) and control inputs u_i (bottom) for the problem in Section 6.2.

ferent optimal control sequence. When the final state is constrained, in general, the problem may turn out to be unfeasible. For example, when the final state $\bar{x}(T) = 0$ is imposed, there is no feasible solution with $T = 12$: a horizon of at least $T = 15$ steps is needed to exactly reach such a target.

7 Conclusions and Future Work

We have proposed a decentralised closed-form iterative formula to solve convex programming problems with a decoupled cost function, linear equality constraints and interval bounds on the decision variables. The algorithm exploits the decentralised communication between decision agents, which are associated with a saturation function and set the value of the decision variables, and information agents, which are associated with integral variables and ensure that the equality constraints are satisfied. Convergence is guaranteed, and is exponential under mild assumptions. Also the discretised version of the algorithm is guaranteed to converge for a small enough step parameter τ.

Several interesting directions are worth exploring. First of all, coupling in the cost function could be considered: for instance, in the case of two variables, a positive coupled quadratic cost would be $f(u_i, u_j) = \alpha u_i^2 + 2\beta u_i u_j + \gamma u_j^2$. However, this would compromise the independence of some control and information agents. After a linear transformation $(u_i, u_j) \to (v_i, v_j)$, providing $f(v_i, v_j) = v_i^2 + v_j^2$, we could handle the problem and extend our result, provided that u_i and u_j are unconstrained (otherwise we would still have coupling in the constraints after the transformation). A further question is whether a more sophisticated discretisation than the Euler scheme can be applied to preserve the sparsity structure and guarantee convergence. An intriguing problem is a possible asynchronous implementation of the scheme, where each agent has its own step parameter τ_i (so that no centralised computation of τ is required). This seems a reasonable possibility, provided that all the decentralised τ_i are small enough; however, we do not have any stability proof so far. These issues are left to future investigation.

Acknowledgements

We thank Alexander Gasnikov for his valuable comments, and the anonymous reviewers for their helpful suggestions. G.G. is supported by the Delft Technology Fellowship Grant at the TU Delft.

References

B. Ataslar and A. Iftar, "A decentralized control approach for transportation networks", *Proc. of the 8th IFAC Symposium on Large Scale Systems*, pp. 348-353, 1998.

F. Bagagiolo, D. Bauso, D. Maggistro and M. Zoppello, "Game theoretic decentralized feedback controls in Markov jump processes", *Journal of Optimization Theory and Applications*, vol. 173, pp. 704–726, 2017.

L. Bakule, “Decentralized control: An overview”, *Annual Reviews in Control*, vol. 32, no. 1, pp. 87–98, 2008.

L. Bakule, “Decentralized control: Status and outlook”, *Annual Reviews in Control*, vol. 38, no. 1, pp. 71–80, 2014.

D. Bauso, F. Blanchini, L. Giarré and R. Pesenti, “The linear saturated control strategy for constrained flow control is asymptotically optimal”, *Automatica*, vol. 49, no. 7, pp. 2206–2212, 2013.

A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, “Foreword: special issue on robust optimization”, *Mathematical Programming*, Ser. B, vol. 107, pp. 1–3, 2006.

D. Bertsimas and A. Thiele, “A Robust Optimization Approach to Inventory Theory”, *Operations Research*, vol. 54, no. 1, pp. 150–168, 2006.

F. Blanchini, S. Miani, “Set–theoretic methods in control”, Birkhauser, Boston, MA, 2015.

F. Blanchini, E. Franco and G. Giordano, “Structured-LMI conditions for stabilizing network-decentralized control”, *Proc. IEEE Conf. on Decision and Control*, Firenze, Italy, pp. 6880–6885, 2013.

F. Blanchini, E. Franco and G. Giordano, “Network-decentralized control strategies for stabilization”, *IEEE Transaction on Automatic Control*, vol. 60, n. 2, pp. 491–496, 2015.

F. Blanchini, E. Franco, G. Giordano, V. Mardanlou, and P. L. Montessoro, “Compartmental flow control: decentralization, robustness and optimality”, *Automatica*, vol. 64, n. 2, pp. 18–28, 2016.

S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, UK, 2004.

S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers”, *Foundations and Trends in Machine Learning*, vol. 3, no. 1, 2011.

M. Breckpot, O. M. Agudelo and B. De Moor, “Flood Control with Model Predictive Control for River Systems with Water Reservoirs”, *Journal of Irrigation and Drainage Engineering*, vol. 139, pp. 532–541, 2013.

M. Breckpot, O. M. Agudelo, P. Meert, P. Willems and B. De Moor, “Flood control of the Demer by using Model Predictive Control”, *Control Engineering Practice*, vol. 21, no. 12, pp. 1776–1787, 2013.

T. H. Chang, A. Nedić and A. Scaglione, “Distributed Constrained Optimization by Consensus-Based Primal-Dual Perturbation Method”, *IEEE Transactions on Automatic Control*, vol. 59, no. 6, 2014.

N. Chatzipanagiotis and M. M. Zavlanos, “A Distributed Algorithm for Convex Constrained Optimization Under Noise”, *IEEE Transactions on Automatic Control*, vol. 61, no. 9, 2016.

A. Cherukuri and J. Cortés, “Distributed algorithms for convex network optimization under non-sparse equality constraints”, *Fifty-fourth Annual Allerton Conference*, Allerton House, UIUC, Illinois, USA September 27–30, 2016.

D. K. Delgoda, S. K. Saleem, M. N. Halgamuge and H. Malano, “Multiple Model Predictive Flood Control in Regulated River Systems with Uncertain Inflows”, Water Resources Management vol. 27, no. 3, pp. 765–790, 2013.

S. Dougherty and M. Guay, “An Extremum-Seeking Controller for Distributed Optimization Over Sensor Networks”, *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 928–933, 2017.

J.C. Duchi, A. Agarwal and M.J. Wainwright, “Dual Averaging for Distributed Optimization: Convergence Analysis and Network Scaling”, *IEEE Transactions on Automatic Control*, vol. 57, no. 3, pp. 592–606, 2002.

A. Falsone, K. Margellos, S. Garatti and M. Prandini, “Dual decomposition for multi-agent distributed optimization with coupling constraints”, *Automatica*, vol.84, pp. 149–158, 2017.

T. Hu and Z. Lin, *Control of systems with actuator saturation*, Birkhauser, Boston, MA, 2001.

A. Iftar, “A linear programming based decentralized routing controller for congested highways”, *Automatica*, vol. 35, no. 2, pp. 279–292, 1999.

A. Iftar and E. J. Davison, “Decentralized control strategies for dynamic routing”, *Optimal Control Applications and Methods*, vol. 23, pp. 329-355, 2002.

B. Johansson, M. Rabi and M. Johansson, “A randomized incremental subgradient method for distributed optimization in networked systems”, *SIAM Journal of Optimization*, vol. 20, no. 3, pp. 1157–1170, 2009.

A. Kozma, C. Conte and M. Diehl, “Benchmarking large-scale distributed convex quadratic programming algorithms”, *Optimization Methods and Software*, 30 (2015): 191–214.

P. Lin, W. Ren and J.A. Farrell, “Distributed Continuous-Time Optimization: Nonuniform Gradient Gains, Finite-Time Convergence, and Convex Constraint Set”, *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2239–2253, 2017.

A. Makhdoumi and A. Ozdaglar, “Convergence Rate of Distributed ADMM Over Networks”, *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5082–5095, 2017.

A. Mokhtari, Q. Ling and A. Ribeiro, “Network Newton Distributed Optimization Methods”, *IEEE Transactions on Signal Processing* vol. 65, no. 1, 2017.

R. A. Montero, D. Schwanenberg, M. Hatz and M. Brinkmann, “Simplified hydraulic modelling in model predictive control of flood mitigation measures along rivers”, *Journal of Applied Water Engineering and Research*, vol. 1, no. 1, pp. 17–27, 2013.

A. Nedić and A. Ozdaglar, “Distributed Subgradient Methods for Multi-Agent Optimization”, *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

A. Nedić, S. Lee and M. Raginsky, “Decentralized Online Optimization with Global Objectives and Local Communication”, *Proc. 2015 American Control Conference*, Chicago, IL, U.S.A., pp. 4497–4503, 2015.

H. Sarimveisa, P. Patrinos, C. D. Tarantilis and C. T. Kiranoudis, “Dynamic modeling and control of supply chain systems: A review”, *Computers & Operations Research* vol. 35, pp. 3530–3561, 2008.

Y. Sekine, “Decentralized Optimization of an Interconnected System”, *IEEE Transactions on Circuit Theory*, vol. 10, no. 2, 1963.

E. A. Silver and R. Peterson, *Decision System for Inventory Management and Production Planning*, Wiley, New York, NY, 1985.

V. Sundarapandian, “An invariance principle for discrete-time nonlinear systems”, *Applied Mathematics Letters*, vol. 16, n. 1, pp. 85–91, 2003.

H. Terelius, U. Topcu and R. M. Murray, “Decentralized Multi-Agent Optimization via Dual Decomposition”, *IFAC World Congress*, vol. 44, no. 1, pp. 11245–11251, 2011.