Distributed Feedback Optimization of Linear Multi-agent Systems

Amir Mehrnoosh and Gianluca Bianchin

ICTEAM Institute, Université catholique de Louvain, Belgium,
amir.mehrnoosh@uclouvain.be
gianluca.bianchin@uclouvain.be

Abstract. Feedback optimization has evolved as an important control approach for optimizing dynamical systems, particularly in terms of steady-state performance objectives. Traditional feedback optimization methods are primarily reliant on centralized systems and controller structures, which suffer from scalability and privacy issues for large-scale applications. This extended abstract investigates a distributed feedback optimization framework in which each agent updates its local control state by averaging inputs from neighboring agents and performing a local negative gradient step. Assuming convexity and smoothness of the cost, we show that this control technique converges to a fixed point. Further, under the weaker assumption of restricted strong convexity of the cost function, we prove that the algorithm achieves linear convergence to a vicinity of the optimal point, with the neighborhood size contingent on the chosen stepsize. Simulation results support the theoretical findings.

Keywords: optimization algorithms, feedback optimization, distributed control

1 Introduction

Optimal steady-state regulation seeks to govern dynamical systems to reach optimal points described by an optimization problem [1]. Traditional techniques split planning and control, with the optimization problem solved offline to determine optimal states that guide controllers. This approach implies that disturbances are known in advance, allowing for accurate optimization. However, in practice, disturbances are frequently unknown, and the primary purpose is to preserve optimality despite these uncertainties. As a result, standard batch optimization strategies fail [2] when disturbances are not precisely known or change during the optimization process.

Recent studies have concentrated on optimal steady-state control problems in a centralized framework. Feedback optimization controllers are well-known for their potential to regulate systems to optimal steady-state conditions while rejecting constant [3] or time-varying disturbances [4], [5]. These controllers convert numerical optimization algorithms into feedback controllers, updating control inputs based on inexact gradients from real-time observations rather than precise plant models or disturbance information. This versatility enables feedback optimization to manage a variety of conditions. However, centralized methods encounter scalability and privacy concerns in large-scale systems. Our work differs from the existing literature by focusing on optimal steady-state control in a distributed, multi-agent architecture, which is consistent with distributed optimization research. Previous work has used distributed gradient descent algorithms [7], [8], and diminishing stepsizes [9] to ensure convergence. Related research focuses on distributed optimization of nonlinear systems [14], which typically separates optimization and tracking control, as well as approaches that use inexact gradients to approximate sensitivity matrices [15]. See also [10]–[13].

This work makes three key contributions. First, we propose a distributed architecture and a control algorithm for optimal steady-state regulation. Our algorithm, inspired by distributed optimization, combines gradient descent with a consensus operation to simultaneously solve the optimization problem and reach an agreement among agents. Second, we prove the convergence of the controller-system state to a fixed point. Our technical analysis provides guidelines for selecting a sufficiently small controller stepsize to ensure system convergence. Third, we establish an explicit bound for the control error. Specifically, under restricted strong convexity assumptions, the controller state converges linearly to a neighborhood of the optimal point. The size of this neighborhood, as shown in previous literature [8], depends on the chosen stepsize. Exact convergence is unattainable because distributed controllers must balance moving towards the minimum gradient and maintaining the agreement with other controllers.

The rest of the document is ordered as follows. Section 2 formalizes the problem of distributed optimal steady-state regulation, while Section 3 shows the proposed controller and gives the work's two primary contributions, the convergence analysis and the proposed controller's error bounds. Section 4 provides numerical simulations to validate our findings, and Section 5 concludes the paper.

2 Preliminaries

In this section, we discuss the notation used in the document. Also, we define the problem and system structure.

Notation. For a symmetric matrix W, we denote its eigenvalues by $\lambda_1(W) \geq \lambda_2(W) \geq \cdots \geq \lambda_n(W)$. We assume that the mixing matrix $W = [w_{ij}]$ is symmetric and doubly stochastic. For $x \in \mathbb{R}^{n_1}$ and $z \in \mathbb{R}^{n_2}$, we denote by $col(x, z) \in \mathbb{R}^{n_1+n_2}$ their column concatenation. The eigenvalues of W are real and sorted in a nonincreasing order $1 = \lambda_1(W) \geq \lambda_2(W) \geq \cdots \geq \lambda_n(W) > -1$. Let the second-largest eigenvalue magnitude of W be:

$$\beta := \max\{|\lambda_2(W)|, |\lambda_n(W)|\}. \tag{1}$$

Problem Setting. We consider a distributed system with N subsystems, $\mathcal{V}_x = \{1, ..., N\}$. The physical couplings between interacting subsystems are described using a graph called system graph (see Fig. 1), $\mathcal{G}_x = (\mathcal{V}_x, \mathcal{E}_x)$ where $\mathcal{E}_x \subseteq \mathcal{V}_x \times \mathcal{V}_x$. Each subsystem $i \in \mathcal{V}$ has a local state $x_i^k \in \mathbb{R}^{n_i}$, that updates for $k \in \mathbb{Z} \geq 0$. Also, we assume that the local controllers have access to global output measurements of the system state as follows,

$$x_i^{k+1} = \sum_{j \in N_i} A_{ij} x_j^k + B_i u_i^k + E_i w_i, \tag{2}$$

$$y^k = \sum_{i=1}^{N} C_i x_i^k + D_i u_i^k, \tag{3}$$

where $\mathcal{N}_i$ is the set of neighbors of i, $u_i^k \in \mathbb{R}^{m_i}$ is the local control input, $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $B_i \in \mathbb{R}^{n_i \times m_i}$, $C_i \in \mathbb{R}^{p \times n_i}$, and $D_i \in \mathbb{R}^{p \times m_i}$. We have $w_i \in \mathbb{R}^{r_i}$ with $E_i \in \mathbb{R}^{n_i \times r_i}$ as a model of constant unknown disturbance acting on subsystem i. We define $n := \sum_i n_i$, $m := \sum_i m_i$,, and $r := \sum_i r_i$.

Equations (2), and (3) can be written in compact form as

$$\begin{aligned} x^{k+1} &= Ax^k + Bu^k + Ew, \\ y^k &= Cx^k + Du^k, \end{aligned} \tag{4}$$

where $x^k = (x_1^k, \ldots, x_N^k) \in \mathbb{R}^n$, $u^k = (u_1^k, \ldots, u_N^k) \in \mathbb{R}^m$, and $w = (w_1, \ldots, w_N) \in \mathbb{R}^r$ are the vectors of states, inputs, and disturbances, respectively. Also, $A = [A_{ij}]$, $B = [B_{ij}]$, $E = [E_{ij}]$, $C = [C_1, \ldots, C_N]$, and $D = [D_1, \ldots, D_N]$, and are block matrices. In what follows, we denote

$$\begin{aligned} G &:= G(1) = C(I - A)^{-1} B + D, \\ H &:= H(1) = C(I - A)^{-1} E, \end{aligned} \tag{5}$$

which are the steady-state transfer functions from u to y and from w to y, respectively.

We assume the following for the system (4).

Assumption 1. *The system (4) is asymptotically stable, i.e., given $Q \succ 0$, we let $P \succ 0$ be such that $A^T P A - P = -Q$. Furthermore, (4) is controllable and observable.* □

If A is not Schur stable, it can be pre-stabilized by standard static state feedback methods. At steady-state, x, u, and y have constant values due to the constant disturbance w as

$$y^{ss} = Gu^{ss} + Hw.$$

This study examines a control problem in which a group of controllers cooperates to optimize the system at steady-state:

$$\min_{u \in \mathbb{R}^m} \sum_{i=1}^{N} \Phi_i(u, Gu + Hw), \tag{6}$$

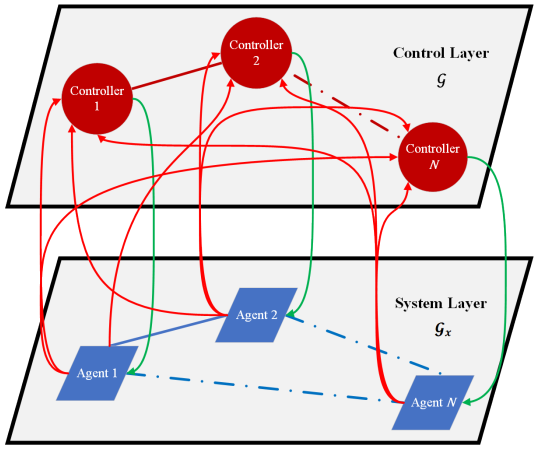


Fig. 1. Distributed system architecture considered in this work (cf. (2)). Each local controller actuates the corresponding subsystem (green lines), by using global feedback information (red lines), see (9).

where Φ_i is supposed to be private and only known to agent i. Solutions to optimization (6) cannot be found explicitly because of the unknown variable w. We consider the following assumptions on the optimization problem.

Assumption 2. *For all i, $(u, y) \mapsto \Phi_i(u, y)$ is proper closed convex, lower bounded, and Lipschitz differentiable with constant L_{Φ_i}.* □

By defining $\Pi^T := [I_m, G^T]$, this assumption guarantees that there exists L_Φ such that

$$||\Pi^T(\nabla\Phi(u, y) - \nabla\Phi(u', y'))|| \leq L_\Phi||col(u, y) - col(u', y')||, \tag{7}$$

for all $y, y' \in \mathbb{R}^n$ and $u, u' \in \mathbb{R}^m$. Also, we define the set of optimizers of (6) by $\mathcal{A}^* := \{(u^*, x^*) : (u^*, x^*)$ *is a first-order optimizer of (6)*$\}$. We assume that this set is closed and nonempty. Also, we denote the compact form of costs as

$$\Phi(u, y) := \sum_{i=1}^{N} \Phi_i(u, y).$$

To represent the collaborative characteristics of the controllers, we will use an undirected graph called the control graph (see Fig. 1) with $\mathcal{G}_u = (\mathcal{V}_u, \mathcal{E}_u)$ where $\mathcal{V}_u = \mathcal{V}_x$ and $\mathcal{E}_u \subseteq \mathcal{V}_u \times \mathcal{V}_u$. To compute a control law cooperatively, two agents must be connected via a communication link in $\mathcal{E}_u$. A graph is connected if there is a path between two nodes. We make the following assumption on the control graph.

Assumption 3. *The graph $\mathcal{G}_u$ is connected. Therefore, there exists weight matrix $W = [w_{ij}] \in \mathbb{R}^{n\times n}$ associated with the communication graph which is a symmetric and doubly stochastic matrix (see (1)).* □

3 Methodology and Analysis

A centralized solution to the steady-state regulation problem (6) has been studied in [6], as well as continuous-time counterparts [3], [4], [5]. In [6], the authors present a gradient-type controller:

$$u^{k+1} = u^k - \eta \nabla \Phi(u^k, y^k), \tag{8}$$

where $\eta > 0$ is the scalar stepsize. The controller (8) uses a gradient-type iteration to solve the optimization (6). It replaces the true gradient $\Pi^T \nabla \Phi(u^k, Gu^k + Hw)$ with a measurement-based version $\Pi^T \nabla \Phi(u^k, y^k)$, eliminating the requirement to measure w.

Implementing (8) needs centralized knowledge of the gradients $\{\nabla \Phi_i\}_{i \in \mathcal{V}_u}$, which is not feasible in our scenario as each Φ_i is known only locally. We propose a distributed version of (8) for our control architecture. Our proposed method involves each agent $i \in \mathcal{V}_u$ maintaining a local copy of $u^k_{(i)} \in \mathbb{R}^m$ and updating it as needed. We have

$$u^{k+1}_{(i)} = \sum_{j=1}^{N} w_{ij} u^k_{(j)} - \eta \Pi^T \nabla \Phi_i(u^k_{(i)}, y^k). \tag{9}$$

To update its local state $u_{(i)}$, each agent i in this control model takes two steps: (i) it computes a weighted average of the states of its neighbors $\sum_{j=1}^{n} w_{ij} u^k_{(j)}$ in order to find a consensus, and (ii) it applies $-\Pi^T \nabla \Phi_i(u^k_{(i)}, y^k)$ to decrease $\Phi_i(u^k_{(i)}, y^k)$. We note that this control law is distributed in the sense that, instead of requiring knowledge of all gradients $\{\nabla \Phi_i\}_{i \in \mathcal{V}_u}$, each agent i simply needs to know the local $\nabla \Phi_i$. Note that each agent needs to know the steady-state map i.e., G, and measure the global output feedback signal y^k. Future work will focus on generalizing the approach to include local output feedback signals.

We apply (9) to control the system (4). We use the following notations of stacked vectors, $u^k_{(1:N)} := col(u^k_{(1)}, u^k_{(2)}, \ldots, u^k_{(N)}) \in \mathbb{R}^{mn}$ for the rest of the document. In vector form, the system (4) controlled by (9) can be written as

$$x^{k+1} = Ax^k + BSu^k_{(1:N)} + Ew, \tag{10}$$

$$y^k = Cx^k + DSu^k_{(1:N)}, \tag{11}$$

$$u^{k+1}_{(i)} = \sum_{j \in \mathcal{N}_i} w_{ij} u^k_{(j)} - \eta \Pi^T \nabla \Phi_i(u^k_{(i)}, y^k), \quad i \in \mathcal{V}_u, \tag{12}$$

where $S \in \mathbb{R}^{m \times mn}$ is given by:

$$S = diag([I_{m_1}, 0, \ldots], [0, I_{m_2}, 0, \ldots], \ldots [0, \ldots, 0, I_{m_n}]).$$

Theorem 1 (Convergence of the state sequence). *Let Assumptions 1-2 hold, W be such that $\beta < 1$, and the stepsize $\eta \leq \bar{\eta} := \min\{\eta_1, \eta_2, \eta_3\}$, where*

$$\begin{aligned} \eta_1 &= \tfrac{1-2\mu+\lambda_n(W)}{L_\Phi}, \\ \eta_2 &= \tfrac{\mu}{\lambda_1(P) L_h^2}, \\ \eta_3 &= \tfrac{\mu \lambda_n(Q)}{\frac{L_\Phi^2}{4} + L_h^2(||A^T P||^2 + \lambda_n(Q)\lambda_1(P)) + L_h L_\Phi ||A^T P||}, \end{aligned} \tag{13}$$

with μ an arbitrary constant, $0 < \mu \leq 1 - \frac{(1-\lambda_n(W))+\eta L_\Phi}{2}$, and $L_h = ||(I - A)^{-1}BS||$. Then, the sequences x^k, $u_{(i)}^k$ generated by (10)-(12) converge. □

We will prove the claim in Tehorem 1 by using La Salle's Invariance Principle in [16]. Theorem 1 demonstrates that with a sufficiently small stepsize η, the sequences of system states x^k and controller states $u_{(i)}^k$ converge asymptotically to a fixed point. This convergence is not straightforward because the proposed controller (10)-(12) incorporates both a consensus step and a gradient step simultaneously. Improperly choosing η can cause oscillations or convergence failure. The upper bounds η_1, η_2, η_3 depend on various system parameters (4) and the optimization problem (6). Notably, the imposed bounds on μ ensure that $\bar{\eta} > 0$ is well-defined.

Furthermore in [16], we derive an error bound on local agent states. Specifically, we show that the local agent states converge geometrically until reaching a neighborhood of the optimal solution. Also, by choosing the proper parameters, reaching to $\mathcal{O}\left(\frac{\eta}{1-\beta}\right)$-neighborhood of the solution set is guaranteed.

4 Simulation Results

In this section, we present our numerical results. We consider a system with three agents, where the A, B, and C matrices are randomly generated from a normal distribution. For A, we generate a network of N agents with $\frac{N(N-1)}{2}\kappa$ edges chosen uniformly at random, where $N = 3$ and $\kappa = 0.5$. We set $n_i = 1$, $\forall i$, so that $n = N$. The control layer network uses the same setting, with the mixing matrix W chosen to be symmetric and doubly stochastic. We ensure that the graph $\mathcal{G}_u$ is connected. We apply our proposed algorithm (10)-(12) to this setup and evaluate its performance.

We apply (10)-(12) to the problem

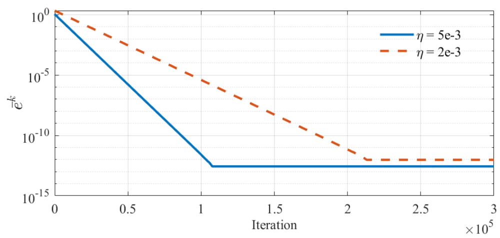
$$\min_{u \in \mathbb{R}^m} \ \frac{1}{2}\left(||u||_R^2 + ||Gu + Hw - y^{ref}||_Q^2\right), \tag{14}$$

where $Q = I_m$, $R = 0.001 I_p$, and $y^{ref} = 0.5 I_p$. Notice that the cost function in (14) is strongly convex in this case.

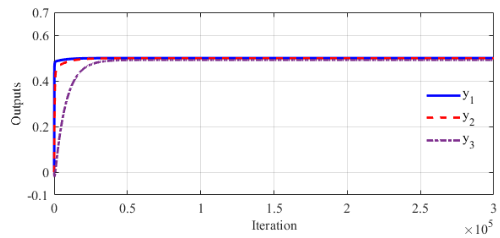
Fig. 2(a) illustrates the convergence of the error $\bar{e}^k$ for two different stepsizes. The error decreases linearly until reaching an $\mathcal{O}(\eta)$-neighborhood, with a smaller η resulting in slower convergence. Fig. 2(b) and 2(c) show that the system outputs reach the desired output, confirming the effectiveness of the proposed algorithm.

5 Conclusions

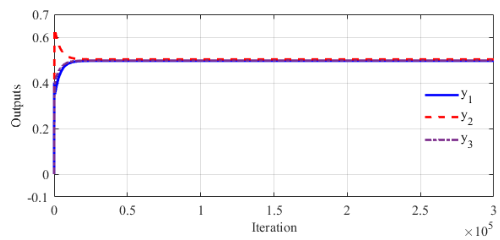
We suggested a distributed controller that can solve optimal steady-state regulation problems while rejecting constant disturbances. This distributed architecture expands effectively with the system size while maintaining the privacy



(a) Comparison of the proposed decentralized algorithm for the problem (14) with different fixed stepsizes.



(b) Outputs of the system with $\eta = 2 \times 10^{-3}$.



(c) Outputs of the system with $\eta = 5 \times 10^{-3}$.

Fig. 2. Error $\bar{e}^k = \frac{1}{N}\sum_{i=1}^{N} u_{(i)}^k - u^{*k}$ and outputs of the proposed decentralized algorithm with different stepsizes, where $u^{*k} = Proj_{\mathcal{A}^*}(\frac{1}{N}\sum_{i=1}^{N} u_{(i)}^k)$

of individual cost functions. Under the convexity and smoothness assumptions, we established that the controller state converges. Under restricted strong convexity assumptions, we showed that the controller converges geometrically to a neighborhood of the optimal solution, consistent with existing distributed optimization literature [14]. Our work paves the way for future research, including scenarios with local output feedback signals, algorithms ensuring exact convergence, the study of constrained optimization objectives, and generalization to nonlinear systems.

References

1. A. Jokic, M. Lazar, and P. van den Bosch, "On constrained steady-state regulation: Dynamic KKT controllers," vol. 54, no. 9, pp. 2250–2254, 2009.
2. E. Dall'Anese, A. Simonetto, S. Becker, and L. Madden, "Optimization and learning with information streams: Time-varying algorithms and applications," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 71–83, 2020.
3. A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Timescale separation in autonomous optimization," vol. 66, no. 2, pp. 611–624, 2021.
4. G. Bianchin, J. Cortes, J. I. Poveda, and E. Dall'Anese, "Time-varying optimization of LTI systems via projected primal-dual gradient flows," vol. 9, no. 1, pp. 474–486, Mar. 2022.
5. G. Bianchin, J. I. Poveda, and E. Dall'Anese, "Online optimization of switched LTI systems using continuous time and hybrid accelerated gradient flows," vol. 146, p. 110579, Aug. 2022.
6. G. Bianchin, M. Vaquero, J. Cortes, and E. Dall'Anese, "Online stochastic optimization for unknown linear systems: Data-driven synthesis and controller analysis," 2024, early Access.
7. A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," 2009, to appear.
8. K. Yuan, Q. Ling, and W. Yin, "On the convergence of decentralized gradient descent," vol. 26, no. 3, pp. 1835–1854, 2016
9. A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," vol. 55, no. 4, pp. 922–938, 2010
10. J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," vol. 57, no. 3, pp. 592–606, 2011.
11. D. Jakovetic, J. Xavier, and J. M. Moura, "Fast distributed gradient methods," vol. 59, no. 5, pp. 1131–1146, 2014.
12. A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi- agent optimization," vol. 54, no. 1, pp. 48–61, 2009.
13. C. Xi, V. S. Mai, R. Xin, E. H. Abed, and U. A. Khan, "Linear convergence in optimization over directed graphs with row-stochastic matrices," vol. 63, no. 10, pp. 3558 3565, 2018.
14. T. Liu, Z. Qin, Y. Hong, and Z.-P. Jiang, "Distributed optimization of nonlinear multiagent systems: A small gain approach," vol. 67, no. 2, pp. 676–691, 2021.
15. W. Wang, Z. He, G. Belgioioso, S. Bolognani, and F. Dörfler, "Decentralized feedback optimization via sensitivity decoupling: Stability and sub-optimality," arXiv preprint, 2023, arXiv:2311.09408.
16. A. Mehrnoosh and G. Bianchin, "Distributed Feedback Optimization of Linear Multi-agent Systems," arXiv preprint, 2024, arXiv:2403.18386.