

# A Parallel Decomposition Scheme for Solving Long-Horizon Optimal Control Problems

Sungho Shin<sup>1</sup>, Timm Faulwasser<sup>2</sup>, Mario Zanon<sup>3</sup>, and Victor M. Zavala<sup>1</sup>

**Abstract**—We present a temporal decomposition scheme for solving long-horizon optimal control problems. In the proposed scheme, the time domain is decomposed into a set of subdomains with partially overlapping regions. Subproblems associated with the subdomains are solved in parallel to obtain local primal-dual trajectories that are assembled to obtain the global trajectories. We provide a sufficient condition that guarantees convergence of the proposed scheme. This condition states that the effect of perturbations on the boundary conditions (i.e., the initial state and terminal dual/adjoint variable) should decay asymptotically as one moves away from the boundaries. This condition also reveals that the scheme converges if the size of the overlap is sufficiently large and that the convergence rate improves with the size of the overlap. We prove that linear quadratic problems satisfy the asymptotic decay condition, and we discuss numerical strategies to determine if the condition holds in more general cases. We draw upon a non-convex optimal control problem to illustrate the performance of the proposed scheme.

## I. INTRODUCTION

Long-horizon optimal control problems (OCPs) arise in model predictive control (MPC) applications such as chemical process systems [1], autonomous vehicle steering [2], and battery systems [3]. They also appear in other application domains such as chemical production planning [4] and electricity production planning [5]. Different decomposition techniques have been reported in the literature to improve computational tractability of these problems including dual decomposition [6], alternating direction method of multipliers [7], dual dynamic programming [8], Gauss-Seidel schemes [9], [10], and parallel Newton schemes [11]. Such decomposition techniques allow scalable solutions of long-horizon OCPs by the use of parallel computers.

In this work, we study the convergence properties of a new *decomposition* paradigm that uses overlapping time domains. This approach is motivated by overlapping Schwartz schemes used for the solution of partial differential equations (PDEs) [12], [13]. In the proposed scheme, the time domain is decomposed into a set of partially overlapping subdomains. Subproblems associated with subdomains are solved in parallel to obtain local primal-dual trajectories by

using current primal-dual information from the neighboring subdomains. The subdomain trajectories are then assembled (this can be interpreted as a projection operator) to update the primal-dual trajectories over the entire domain, and the procedure is repeated. We provide a sufficient condition that guarantees convergence of the proposed decomposition algorithm when applied to OCPs. This condition applies to linear and nonlinear problems. Specifically, the condition indicates that convergence of the decomposition scheme is guaranteed provided that the sensitivity of the primal-dual (state-adjoint) trajectories to perturbations in initial states and terminal cost gradients decay asymptotically as one moves away from the boundaries. We call this condition *asymptotic decay of sensitivity* (ADS). The condition also reveals that the algorithm converges provided that the size of the overlap is sufficiently large and that the convergence rate improves as the size of the overlapping regions increases.

ADS-like properties have been recently explored in the literature. Xu et. al. recently showed that an ADS condition (in the primal space) holds under a time-varying and inequality-constrained linear quadratic (LQ) control setting [15], [16]. To prove this, the authors assumed uniformly complete controllability and exploited the algebraic structure of the Riccati equation. The authors used the primal ADS property to show that trajectories of an overlapping temporal decomposition scheme approximate those of the long-horizon problem and that the approximation error decays as the size of the overlap increases. The authors also showed that receding horizon control provides approximate trajectories and that the error converges as the size of the overlapping regions increase. These works do not provide an algorithmic scheme that delivers optimal trajectories (for a given size of the overlapping region), as we do in this work. Shin et. al. recently established a primal ADS condition for general graph-structured, unconstrained quadratic programs and showed that this condition guarantees the convergence of an iterative overlapping decomposition algorithm [14]. The condition established in this work is specialized to constrained OCPs and operates in the primal-dual space. A primal ADS property has also been established in [17, Lemma 5]. This condition is exploited by the authors to establish the stability of economic MPC.

The paper is organized as follows. In Section II, we present the basic setting and describe the proposed decomposition scheme. In Section III, we propose a primal-dual ADS condition that guarantees convergence of the algorithm. In Section IV, we show that ADS holds for a simplified LQ setting. In Section V, we demonstrate the proposed scheme

<sup>1</sup>S. Shin and V. M. Zavala are with the Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI 53706 USA (e-mail: sungho.shin@wisc.edu; victor.zavala@wisc.edu).

<sup>2</sup>Timm Faulwasser is with Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany. (e-mail: timm.faulwasser@ieee.org)

<sup>3</sup>Mario Zanon is with IMT Lucca, 55100 Lucca, Italy.(e-mail: mario.zanon@imtlucca.it)

TF acknowledges support by the Bundesministerium für Bildung und Forschung (BMBF), Grant 05M18CKA.

using a nonlinear economic MPC problem.

## II. BASIC DEFINITIONS AND SETTING

We consider an OCP with a time domain set  $\mathbb{I}_{M:N}$ , an initial state  $x_M^* \in \mathbb{R}^{n_x}$ , and a terminal cost gradient  $\lambda_N^* \in \mathbb{R}^{n_x}$  of the form:

$$\min_{\substack{x_{M:N} \\ u_{M:N-1}}} \sum_{i=M}^{N-1} \ell(x_i, u_i) + (\lambda_N^*)^\top x_N \quad (1a)$$

$$\text{s.t. } x_M = x_M^* \quad (\lambda_M) \quad (1b)$$

$$x_i = f(x_{i-1}, u_{i-1}) \quad (\lambda_i) \quad \forall i \in \mathbb{I}_{M+1:N} \quad (1c)$$

$$g(x_i, u_i) \leq 0 \quad (\mu_i), \quad \forall i \in \mathbb{I}_{M:N-1} \quad (1d)$$

We denote this problem as  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$ . Here,  $x_i \in \mathbb{R}^{n_x}$  and  $u_i \in \mathbb{R}^{n_u}$  are the state and input variables at time  $i$ ;  $\lambda_i \in \mathbb{R}^{n_x}$  and  $\mu_i \in \mathbb{R}^{n_g}$  are the dual variables associated with (1b)-(1c) and (1d), respectively. Symbol  $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_x}$  is the dynamic mapping,  $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$  is the stage cost mapping, and  $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_g}$  is the constraint mapping. We denote the set of real numbers and the set of integers as  $\mathbb{R}$  and  $\mathbb{Z}$ , respectively, and we define  $\mathbb{I}_{M:N} := \mathbb{Z} \cap [M, N]$ . We use the syntax  $(x_1, x_2, \dots, x_n) := [x_1^\top \ x_2^\top \ \dots \ x_n^\top]^\top$  and  $x_{M:N} := (x_M, x_{M+1}, \dots, x_N)$ . We denote primal-dual pairs as  $z_i := (x_i, \lambda_i)$ .

*Remark 1:* From the KKT conditions of (1), it follows that  $\lambda_N = \lambda_N^*$ . Therefore, incorporating the terminal penalty term  $(\lambda_N^*)^\top x_N$  in the objective function as (1a) essentially constitute the terminal constraint of the dual variable  $\lambda$  at  $i = N$ .

We now describe the proposed decomposition scheme for solving  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$ . We partition the time domain  $\mathbb{I}_{M:N}$  into a collection of  $K$  non-overlapping (i.e., disjoint) sets of the form  $\mathbb{I}_{M_1:N_1}, \dots, \mathbb{I}_{M_K:N_K}$ . We also define a collection of overlapping sets  $\mathbb{I}_{M_1^\omega:N_1^\omega}, \dots, \mathbb{I}_{M_K^\omega:N_K^\omega}$  satisfying

$$M_k^\omega = \max(M_k - \omega, M), \quad N_k^\omega = \min(N_k + \omega, N)$$

for any  $k \in \mathbb{I}_{1:K}$ . We call  $\omega \in \mathbb{Z}_{>0}$  the *size of the overlap*. The non-overlapping and overlapping time domains with  $\omega = 1$  are illustrated in Figure 1.

The scheme starts with an initial guess of the primal  $x_{M:N}^{(0)}$  and dual  $\lambda_{M:N}^{(0)}$  trajectories with  $x_M^{(0)} = x_M^*$  and  $\lambda_N^{(0)} = \lambda_N^*$ . For each subdomain index  $k \in \mathbb{I}_{1:K}$  and iteration counter  $t \in \mathbb{Z}_{>0}$ , subproblems  $\mathcal{P}_{M_k^\omega:N_k^\omega}(x_{M_k^\omega}^{(t)}, \lambda_{N_k^\omega}^{(t)})$  are solved in parallel. Each subproblem requires current primal-dual information from the neighboring subproblems (this information enters in the initial condition and terminal penalty). The solution of subproblem  $\mathcal{P}_{M_k^\omega:N_k^\omega}(x_{M_k^\omega}^{(t)}, \lambda_{N_k^\omega}^{(t)})$  yields the local primal-dual trajectories  $\hat{z}_{M_k^\omega:N_k^\omega}^{(t+1,k)}$ . These trajectories are restricted to the non-overlapping subdomains  $\mathbb{I}_{M_k:N_k}$  by using the restriction  $z_{M_k:N_k}^{(t+1)} := \hat{z}_{M_k^\omega:N_k^\omega}^{(t+1,k)}$ . The solutions at  $\mathbb{I}_{M_k^\omega:N_k^\omega} \setminus \mathbb{I}_{M_k:N_k}$  are discarded. This restriction procedure over  $k \in \mathbb{I}_{1:K}$  can be seen as a projection that assembles the entire primal-dual trajectory  $z_{M:N}^{(t+1)}$  that is in turn used as the next guess in the algorithm.

The trajectory assembling procedure is sketched in Figure 1. The proposed decomposition scheme is summarized in Algorithm 1 and can be stated compactly as follows:

$$z_{M_k:N_k}^{(t+1)} \leftarrow \mathcal{P}_{M_k^\omega:N_k^\omega}(x_{M_k^\omega}^{(t)}, \lambda_{N_k^\omega}^{(t)}), \quad k \in \mathbb{I}_{1:K}, \quad t \in \mathbb{Z}_{>0}. \quad (2)$$

---

### Algorithm 1 Decomposition Scheme with Overlap

---

Initialize  $z_{M:N}^{(0)}$  and  $t \leftarrow 0$   
**while** termination criteria unsatisfied **do**  
  **for** (in parallel)  $k = 1$  to  $K$  **do**  
    Solve  $\mathcal{P}_{M_k^\omega:N_k^\omega}(x_{M_k^\omega}^{(t)}, \lambda_{N_k^\omega}^{(t)})$  to obtain  $\hat{z}_{M_k^\omega:N_k^\omega}^{(t+1,k)}$   
    Apply restriction  $z_{M_k:N_k}^{(t+1)} := \hat{z}_{M_k^\omega:N_k^\omega}^{(t+1,k)}$   
  **end for**  
  Evaluate termination criteria  
   $t \leftarrow t + 1$   
**end while**

---

## III. CONVERGENCE RESULTS

We now provide a sufficient condition for (1) that guarantees convergence of the decomposition algorithm (2). We begin by making the following existence and uniqueness assumption on solutions of problem (1).

*Assumption 1 (Existence and uniqueness of solution):*

There exist sets  $X$  and  $\Lambda \subseteq \mathbb{R}^{n_x}$  such that, for any  $M, N \in \mathbb{Z}_{>0}$  with  $M < N$  and  $(x_M^*, \lambda_N^*) \in X \times \Lambda$ , there exists a unique primal-dual solution  $z_{M:N}^*$  of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  with  $z_i \in X \times \Lambda$  and for any  $i \in \mathbb{I}_{M:N}$ .

*Remark 2:* The solution of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  always satisfies  $x_M = x_M^*$  (by (1b)) and  $\lambda_N = \lambda_N^*$  (see Remark 1), and thus we can write the solution of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  as  $z_{M:N}^* = (x_{M:N}^*, \lambda_{M:N}^*)$ . Because the scheme (2) is mostly concerned with the state and the adjoints (this is the information exchanged between domains), we will not explicitly indicate the controls  $u_{M:N-1}^*$  and multipliers  $\mu_{M:N-1}^*$  in the nomenclature.

We now state the following principle of optimality result:

*Lemma 1:* Assume that problem (1) satisfies Assumption 1 and consider  $P, M, N, L \in \mathbb{Z}_{>0}$  with  $P \leq M < N \leq L$  and  $(x_P^*, \lambda_L^*) \in X \times \Lambda$ . Let  $z_{P:L}^*$  be the solution of  $\mathcal{P}_{P:L}(x_P^*, \lambda_L^*)$ , then  $z_{M:N}^*$  is the solution of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$ .

*Proof:* We define the Lagrangian  $\mathcal{L}_{M:N}(\cdot; \cdot)$  of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  as:

$$\begin{aligned} \mathcal{L}_{M:N}(x_{M:N}, u_{M:N-1}, \lambda_{M:N}, \mu_{M:N-1}; x_M^*, \lambda_N^*) \quad (3) \\ := \sum_{i=M}^{N-1} \ell(x_i, u_i) + \lambda_M^\top (x_M - x_M^*) + (\lambda_N^*)^\top x_N \\ + \sum_{i=M+1}^N \lambda_i^\top (x_i - f(x_{i-1}, u_{i-1})) + \sum_{i=M}^{N-1} \mu_i^\top g(x_i, u_i) \end{aligned}$$

From Assumption 1, the solution of  $\mathcal{P}_{1:L}(x_1^*, \lambda_L^*)$  exists and is unique and this implies that the KKT conditions hold at

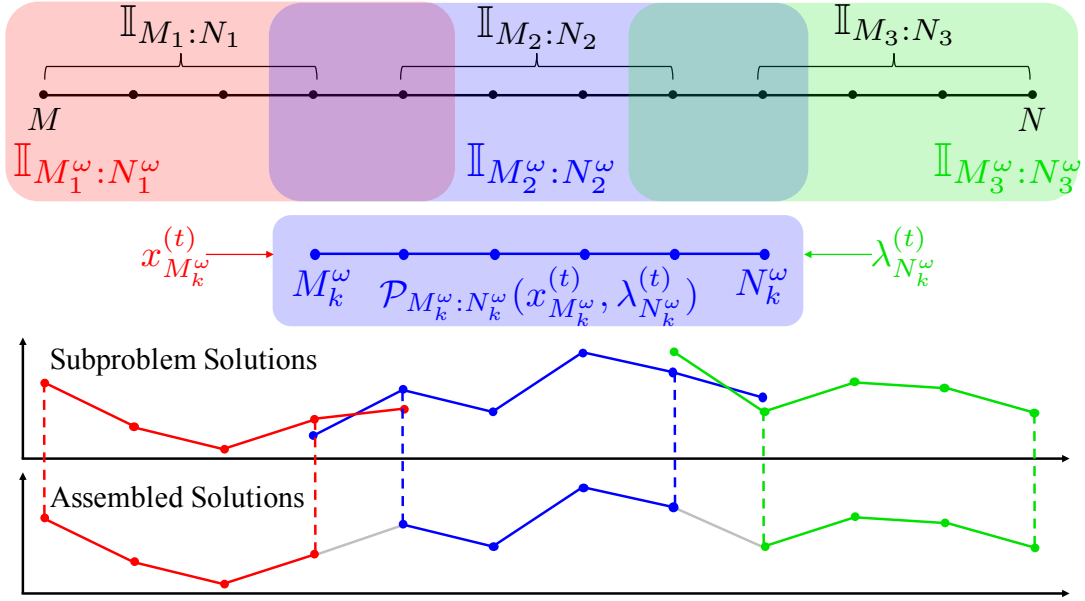


Fig. 1. Sketch of parallel decomposition algorithm.

$x_{M:N}^*, u_{M:N-1}^*, \lambda_{M:N}^*, \mu_{M:N-1}^*$ ; that is,

$$\lambda_i^* + (\nabla_x f_i^*)^\top \lambda_{i+1}^* + \nabla_x \ell_i^* + (\nabla_x g_i^*)^\top \mu_i^* = 0 \quad (4a)$$

$$(\nabla_u f_i^*)^\top \lambda_{i+1}^* + \nabla_u \ell_i^* + (\nabla_u g_i^*)^\top \mu_i^* = 0 \quad (4b)$$

$$x_{i+1}^* = f_i^* \quad (4c)$$

$$g_i^* \leq 0, \quad \mu_i^* \geq 0, \quad \text{diag}(\mu_i^*) g_i^* = 0 \quad (4d)$$

holds for any  $i \in \mathbb{I}_{M:N-1}$ , where  $f_i^* := f(x_i^*, u_i^*)$ ,  $\nabla f_i^* := \nabla f(x_i^*, u_i^*)$ ,  $\nabla g_i^* := \nabla g(x_i^*, u_i^*)$ ,  $\nabla \ell_i^* := \nabla \ell(x_i^*, u_i^*)$ .

From (4), we observe that the KKT conditions of  $\mathcal{P}_{P:L}(x_1^*, \lambda_L^*)$  with  $(x_{P:L}^*, u_{P:L-1}^*, \lambda_{P:L}^*, \mu_{P:L-1}^*)$  imply that the KKT conditions of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  hold with  $(x_{M:N}^*, u_{M:N-1}^*, \lambda_{M:N}^*, \mu_{M:N-1}^*)$ . This consistency is the result of using the duals as a terminal penalty in (1).

By Assumption 1, we have that  $(x_M^*, \lambda_N^*) \in X \times \Lambda$ . This implies that the solution of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  exists and is unique, and thus the KKT point is the unique solution. Thus,  $x_{M:N-1}^*, u_{M:N-1}^*, \lambda_{M:N}^*$ , and  $\mu_{M:N-1}^*$  form the solutions of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$ . ■

Lemma 1 justifies the structure of (1); specifically, primal and dual information of the neighboring subproblems should be incorporated as initial states and terminal penalties, respectively. Furthermore, it implies that the primal-dual trajectory of the subproblems can be assembled to obtain the optimal primal-dual trajectory of the entire problem if the boundary conditions are set to the optimal values. From this we also see that a receding-horizon control scheme delivers the solution of the long-horizon problem if the terminal cost gradients are obtained from the dual solution of the long-horizon problem. Recently the effect of terminal cost gradient on the solution trajectory of optimal control problems and its connection with dual variables has been investigated in [17], [18]. In particular, it is shown that economic MPC can achieve asymptotic stability without terminal constraints by

incorporating the dual of the underlying steady-state problem as a terminal cost gradient. Furthermore, it is shown that LQ economic MPC can be stabilized by adaptively tuning the terminal cost gradient. Such observations align with our observations of Lemma 1. Specifically, the quality of the solutions of OCPs (which in general improves with the horizon length) can also be improved by choosing a suitable terminal cost gradient.

We now state our sufficiency condition for convergence, that we call asymptotic decay of sensitivity (ADS).

*Property 1 (Asymptotic decay of sensitivity):* Assume that (1) satisfies Assumption 1. For given  $M, N \in \mathbb{Z}_{>0}$  with  $M < N$  and  $(x_M^*, \lambda_N^*)$ ,  $(\tilde{x}_M, \tilde{\lambda}_N) \in X \times \Lambda$ , let  $z_{M:N}^*$  and  $\tilde{z}_{M:N}$  be the solutions of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  and  $\mathcal{P}_{M:N}(\tilde{x}_M, \tilde{\lambda}_N)$ , respectively. There exist  $\{\varepsilon_i\}_{i \in \mathbb{Z}_{>0}}$  with  $\varepsilon_i \rightarrow 0$  as  $i \rightarrow \infty$  such that:

$$\|z_i^* - \tilde{z}_i\|_\infty \leq \varepsilon_{i-M} \|x_M^* - \tilde{x}_M\|_\infty + \varepsilon_{N-i} \|\lambda_N^* - \tilde{\lambda}_N\|_\infty$$

holds for any  $i \in \mathbb{I}_{M:N}$ .

Property 1 implies that the solution  $z_i^* = (x_i^*, \lambda_i^*)$  of  $\mathcal{P}_{M:N}(x_M^*, \lambda_N^*)$  at time  $i$  becomes less sensitive to perturbations in the initial state  $x_M^*$  and the terminal cost gradient  $\lambda_N^*$  as the time index  $i$  moves away from the boundary.

*Remark 3:* It is important that the sequence  $\{\varepsilon_i\}_{i \in \mathbb{Z}_{>0}}$  is a *uniform parameter* that does not depend on  $M$  and  $N$  and that does not depend on the choice of the boundary conditions (i.e.,  $x_M^*$  and  $\lambda_N^*$ ). This enables a uniform bound that holds for different subproblems.

*Remark 4 (Relation with turnpike properties):* Property 1 is related but not equivalent to so-called turnpike properties of OCPs (see [19], [20]). Property 1 establishes a relationship between two solution trajectories, while a classical turnpike property compares a single solution trajectory with the steady-state solution. In addition, Property 1 requires





Now we show that  $\bar{x}_1, \dots, \bar{x}_{n_x}$  are linearly independent. To establish a contradiction, suppose that  $\bar{x}_1, \dots, \bar{x}_{n_x}$  are linearly dependent; then there exist not all-zero  $\alpha_1, \dots, \alpha_{n_x}$  such that  $\alpha_1 \bar{x}_1 + \dots + \alpha_{n_x} \bar{x}_{n_x} = 0$  holds. This implies that

$$\alpha_1 (bu_0) + \dots + \alpha_{n_x} \left( \sum_{i=0}^{n_x-1} A^{n_x-1-i} bu_i \right) = 0.$$

By rearranging, we obtain

$$(\alpha_1 u_0 + \dots + \alpha_{n_x} u_{n_x-1}) b + \dots + (\alpha_{n_x} u_0) A^{n_x-1} b = 0. \quad (21)$$

Since  $u_0 = 1$  and  $\alpha_1, \dots, \alpha_{n_x}$  are not all-zero, a non-zero coefficient exists in (21), and thus  $b, \dots, A^{n_x-1} b$  are linearly dependent. This contradicts the assumption that  $(A, b)$  is controllable. Consequently,  $\bar{x}_1, \dots, \bar{x}_{n_x}$  are linearly independent.

From the linear independence of  $\bar{x}_1, \dots, \bar{x}_{n_x}$ , one can show that  $\hat{Z}$  has full column rank and thus  $\hat{H}$  is positive definite. From (20) one can show that

$$\begin{aligned} p^\top (Z^\top H Z) p &\geq \sum_{i=1}^{N-1} \lambda_{\min}(\hat{H}) \|p_{i-N+1:i}\|_2^2 \\ &\geq \lambda_{\min}(\hat{H}) \|p\|_2^2 = \lambda_{\min}(\hat{H}). \end{aligned}$$

Observe that  $\hat{H}$  is independent of the length of the problem. Thus, we can set uniform parameter  $\lambda_1 := \lambda_{\min}(\hat{H})$ . ■

By Proposition 1 and Lemma 2, we have that

$$\left| (\bar{H})_{i,j} \right| \leq \frac{1}{\lambda_1} \left( \frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \right)^{|i-j|/n_x}, \quad \forall i, j \in \mathbb{I}_{1:N-1} \quad (23)$$

where  $\lambda_1$  and  $\lambda_2$  are the uniform upper and lower bound of the eigenvalues of  $Z^\top H Z$  established in Lemma 2.

Finally, inspecting (18) we observe that  $Z$  and  $W$  are sparse. In particular, there exists uniform parameter  $N_s$  such that the following holds for  $|j-i| \geq N_s$ .

$$(Z)_{\alpha(i),j} = 0, \quad (W)_{\beta(i),j} = 0 \quad (24)$$

where  $\alpha(i) := \mathbb{I}_{(i-1)(n_x+1)+1:i n_x+i-1}$  and  $\beta(i) := \mathbb{I}_{(i-1)n_x+1:i n_x}$  (these index sets correspond to the index of  $x_i$  and  $\lambda_i$  among the entries of  $w$  and  $\lambda$ , respectively). Using the sparsity structure identified in (24), one can show the following quantities are less than or equal to  $L^2 \|(\bar{H})_{i \pm N_s, j \pm N_s}\|_\infty$

$$\|(Z \bar{H} Z^\top)_{\alpha(i), \alpha(j)}\|_\infty, \|(W \bar{H} Z^\top)_{\beta(i), \alpha(j)}\|_\infty \quad (25a)$$

$$\|(Z \bar{H} W^\top)_{\alpha(i), \beta(j)}\|_\infty, \|(W \bar{H} W^\top)_{\beta(i), \beta(j)}\|_\infty \quad (25b)$$

where we use the syntax  $i \pm N_s := \mathbb{I}_{i-N_s:i+N_s}$  and define  $(\bar{H})_{i,j} := 0$  if  $\{i,j\} \not\subseteq \mathbb{I}_{1:N-1}$  for convenience. From (18) and (25), we have

$$\begin{aligned} \|x_i^* - \tilde{x}_i\|_\infty &\leq L^2 \|(\bar{H})_{i \pm N_s, N-1 \pm N_s}\|_\infty \|\Delta \lambda_N\|_\infty \\ &\quad + L^2 \|(\bar{H})_{i \pm N_s, 1 \pm N_s}\|_\infty \|\Delta x_1\|_\infty \\ \|\lambda_i^* - \tilde{\lambda}_i\|_\infty &\leq L^2 \|(\bar{H})_{i \pm N_s, 1 \pm N_s}\|_\infty \|\Delta x_1\|_\infty \\ &\quad + (L^2 \|(\bar{H})_{i \pm N_s, N-1 \pm N_s}\|_\infty + \mathbf{1}_{i=N}) \|\Delta \lambda_N\|_\infty. \end{aligned}$$

By (23) and  $\mathbf{1}_{i=N} \leq \rho^{N-i-2N_s-1}$ , we have that

$$\begin{aligned} \|x_i^* - \tilde{x}_i\|_\infty &\leq L^2 \frac{2N_s+1}{\lambda_1} \rho^{N-i-2N_s-1} \|\Delta \lambda_N\|_\infty \\ &\quad + L^2 \frac{2N_s+1}{\lambda_1} \rho^{i-2N_s-1} \|\Delta x_1\|_\infty \end{aligned} \quad (27a)$$

$$\begin{aligned} \|\lambda_i^* - \tilde{\lambda}_i\|_\infty &\leq L^2 \frac{2N_s+1}{\lambda_1} \rho^{i-2N_s-1} \|\Delta x_1\|_\infty \\ &\quad + (L^2 \frac{2N_s+1}{\lambda_1} + 1) \rho^{N-i-2N_s-1} \|\Delta \lambda_N\|_\infty \end{aligned} \quad (27b)$$

for  $i \in \mathbb{I}_{M:N}$  and  $\rho := (\lambda_2 - \lambda_1)/(\lambda_2 + \lambda_1)$ . We define

$$\varepsilon_i := \left( L^2 \frac{2N_s+1}{\lambda_1} + 1 \right) \rho^{i-2N_s-1}, \quad \forall i \in \mathbb{Z}_{>0}. \quad (28)$$

Note that  $\{\varepsilon_i\}_{i \in \mathbb{Z}_{>0}}$  is a uniform parameter and (27)-(28) establish Property (1). This concludes the proof. ■

Theorem 2 implies that the reduced Hessian is positive definite—a key requirement in optimization sensitivity results as the solution must be locally unique and bounded perturbations yield bounded differences in solutions. We thus expect that Theorem 2 can be generalized to LQ OCPs with multiple inputs, time-variant objectives and dynamics, and inequality constraints by exploiting algebraic properties. Here, we focus on a simple setting due to space limitations and to keep the presentation clear. For an even more general setting (nonlinear, inequality-constrained) setting, one can seek to validate the sensitivity property numerically by using simulations. We show how to do this in the next section.

## V. NUMERICAL EXAMPLES

We use a nonlinear OCP to illustrate that Property 1 guarantees convergence of the decomposition algorithm and to highlight that the approach achieves faster solutions than off-the-shelf solvers. We consider a nonlinear economic MPC problem for a chemical reactor [20], [21]. The dynamics of the reactor are given by:

$$\frac{dc_A}{dt} = 1 - 10^4 (c_A)^2 e^{-\frac{1}{T}} + 400 c_A e^{-\frac{0.55}{T}} - c_A \quad (29a)$$

$$\frac{dc_B}{dt} = 10^4 (c_A)^2 e^{-\frac{1}{T}} - c_B \quad (29b)$$

$$\frac{dT}{dt} = u - T \quad (29c)$$

Here,  $c_A$ ,  $c_B$ , and  $T$  are the concentration of  $A$ , the concentration of  $B$ , and the temperature;  $x := (c_A, c_B, T)$  are the state variables;  $u$  is the input variable. The objective is a combination of an economic objective (maximizing the production of  $B$ ) and a convex regularization term:

$$\ell(x, u) := -c_B + \rho_{\text{reg}} (u - u_s)^2 \quad (29d)$$

where  $u_s$  is determined by solving the underlying steady-state optimization problem that minimizes  $-c_B$  (see [19]). The following inequality constraints are enforced on the states and the inputs:

$$c_A, c_B, T \geq 0 \quad \text{and} \quad 0.049 \leq u \leq 0.449 \quad (29e)$$

The control step length is 1 sec and the differential equations are discretized using an implicit Euler scheme with a step

length of 0.25 sec. We implemented a parallel version of the scheme in Julia. Problems were formulated in the modeling language JuMP [22], and were solved with the nonlinear programming solver IPOPT [23]. The scheme was executed on an Intel Xeon CPU E5-2698 v3 processor running at 2.30GHz.

We use numerical simulations to verify that Property 1 holds. Here, we assess the sensitivity of the primal-dual trajectories against perturbations to the initial state and the terminal cost gradient. The reference problem is formulated with  $N = 600$  (i.e., 10 mins) and boundary conditions with  $(x_1^*, \lambda_N^*)$ . The boundary conditions are perturbed as  $(x_1^*, \lambda_N^*) + \delta$ , where  $\delta$  is sampled from a normal random variable. It is known that the OCP with (29) has a periodic optimal solution when  $\rho_{\text{reg}} = 0$  and a steady-state solution with  $\rho_{\text{reg}} = 0.5$  [20]. The solutions of the reference (unperturbed) problem and 30 samples of perturbed problems are shown in Figure 2 ( $\rho_{\text{reg}} = 0.5$ ) and Figure 3 ( $\rho_{\text{reg}} = 0$ ). When  $\rho_{\text{reg}} = 0.5$  we see that, for both primal and dual solutions, the distances from the reference trajectories are small in the center of the time domain and grow as we approach the boundary. On the other hand, when  $\rho_{\text{reg}} = 0$ , such convergence is not observed. This indicates that Property 1 holds for  $\rho_{\text{reg}} = 0.5$  but does not hold for  $\rho_{\text{reg}} = 0$ . This implies that the ADS property strongly depends on the stage cost  $\ell$ . For high-dimensional systems where it is difficult to graphically assess the ADS property, one may consider assessing the error trajectory based on the norm of deviation from the reference trajectory.

We next verify that the decomposition scheme converges and that the convergence rate improves with the size of the overlap. A problem with  $N = 4,800$  (i.e., 80 mins) was decomposed into 8 subdomains and solved with  $\omega = 2, 4$ , and 8 and  $\rho_{\text{reg}} = 0.5$ . Figure 4 (left) shows the evolution of the residuals and confirms that the rate improves with the overlap. The algorithm did not converge when  $\rho_{\text{reg}} = 0$ . These observations reinforce the key role of Property 1.

We also explore the computational efficiency achieved via parallel decomposition. A problem with  $N = 4,800$  was solved using IPOPT and we compared the solution time against that of the decomposition scheme with an overlap of  $\omega = 8$  and  $\rho_{\text{reg}} = 0.5$ . The solution found by the decomposition scheme was equal to the solution from IPOPT. Figure 4 (right) clearly illustrates that the solution time decreases as the number of computing cores increases (this also increases the number of subdomains). We can see that, if a sufficient number of cores are used, the proposed scheme becomes faster than IPOPT.

## VI. CONCLUSIONS

We have presented a temporal decomposition scheme for long-horizon OCPs that solves subproblems on overlapping time domains. We have proposed an asymptotic sensitivity decay property that guarantees convergence of the algorithm. This property indicates that the sensitivity of the primal and dual trajectories to perturbations on the boundary conditions

(initial state and terminal cost gradient) decays asymptotically as one move away from the boundaries. This property also indicates that the scheme converges if the overlap is sufficiently large and the convergence rate improves with the size of the overlap. We have demonstrated that the solution time of long-horizon OCPs can be improved with the proposed decomposition method.

## REFERENCES

- [1] M. Baldea and P. Daoutidis, "Control of integrated process networks—a multi-time scale perspective," *Computers & chemical engineering*, vol. 31, no. 5-6, pp. 426–444, 2007.
- [2] P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, and D. Hrovat, "Predictive active steering control for autonomous vehicle systems," *IEEE Transactions on control systems technology*, vol. 15, no. 3, pp. 566–580, 2007.
- [3] R. Kumar, M. J. Wenzel, M. J. Ellis, M. N. ElBsat, K. H. Drees, and V. M. Zavala, "Hierarchical mpc schemes for periodic systems using stochastic programming," *arXiv preprint arXiv:1804.10866*, 2018.
- [4] J. R. Jackson and I. E. Grossmann, "Temporal decomposition scheme for nonlinear multisite production planning and distribution models," *Industrial & engineering chemistry research*, vol. 42, no. 13, pp. 3045–3055, 2003.
- [5] C. Barrows, M. Hummon, W. Jones, and E. Hale, "Time domain partitioning of electricity production cost simulations," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2014.
- [6] P. Giselsson, M. D. Doan, T. Keviczky, B. De Schutter, and A. Rantzer, "Accelerated gradient methods and dual decomposition in distributed model predictive control," *Automatica*, vol. 49, no. 3, pp. 829–833, 2013.
- [7] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [8] A. M. Geoffrion, "Generalized benders decomposition," *Journal of optimization theory and applications*, vol. 10, no. 4, pp. 237–260, 1972.
- [9] V. M. Zavala, "New architectures for hierarchical predictive control," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 43–48, 2016.
- [10] S. Shin and V. M. Zavala, "Multi-grid schemes for multi-scale coordination of energy systems," in *Energy Markets and Responsive Grids*. Springer, 2018, pp. 195–222.
- [11] H. Deng and T. Ohtsuka, "A parallel code generation toolkit for nonlinear model predictive control," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 4920–4926.
- [12] T. Mathew, *Domain decomposition methods for the numerical solution of partial differential equations*. Springer Science & Business Media, 2008, vol. 61.
- [13] M. Dryja and O. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*. Ultracomputer Research Laboratory, Univ., Courant Inst. of Mathematical Sciences, Division of Computer Science, 1987.
- [14] S. Shin, V. M. Zavala, and M. Anitescu, "Decentralized schemes with overlap for solving graph-structured optimization problems," *arXiv preprint arXiv:1810.00491*, 2018.
- [15] W. Xu and M. Anitescu, "Exponentially accurate temporal decomposition for long-horizon linear-quadratic dynamic optimization," *SIAM Journal on Optimization*, vol. 28, no. 3, pp. 2541–2573, 2018.
- [16] —, "Exponentially convergent receding horizon strategy for constrained optimal control," 2017, preprint ANL/MCS-P9015-1017.
- [17] T. Faulwasser and M. Zanon, "Asymptotic stability of economic NMPC: The importance of adjoints," *IFAC-PapersOnLine*, vol. 51, no. 20, pp. 157–168, 2018, 6th IFAC Conference on Nonlinear Model Predictive Control (2018 NMPC).
- [18] M. Zanon and T. Faulwasser, "Economic MPC without terminal constraints: Gradient-correcting end penalties enforce stability," *Journal of Process Control*, vol. 63, pp. 1–14, 3 2018.
- [19] T. Faulwasser, M. Korda, C. Jones, and D. Bonvin, "On turnpike and dissipativity properties of continuous-time optimal control problems," *Automatica*, vol. 81, pp. 297–304, April 2017.
- [20] T. Faulwasser, L. Grüne, M. A. Müller, *et al.*, "Economic nonlinear model predictive control," *Foundations and Trends® in Systems and Control*, vol. 5, no. 1, pp. 1–98, 2018.

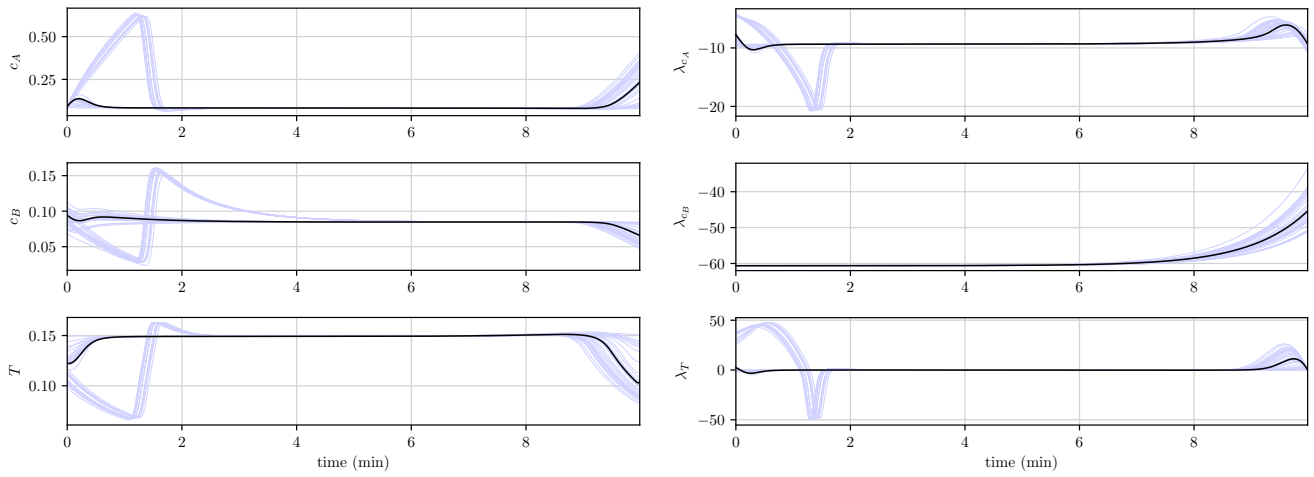


Fig. 2. The primal and dual trajectories of the reference problem (black) and 30 perturbed problems (light blue) with  $N = 180$  and  $\rho_{\text{reg}} = 0.5$ .

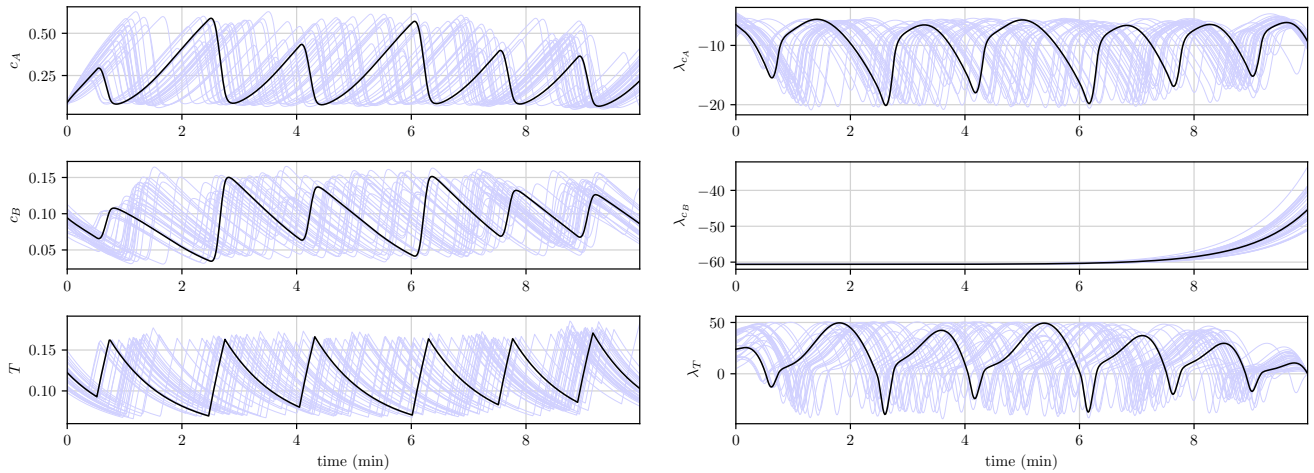


Fig. 3. The primal and dual trajectories of the reference problem (black) and 30 perturbed problems (light blue) with  $N = 180$  and  $\rho_{\text{reg}} = 0$ .

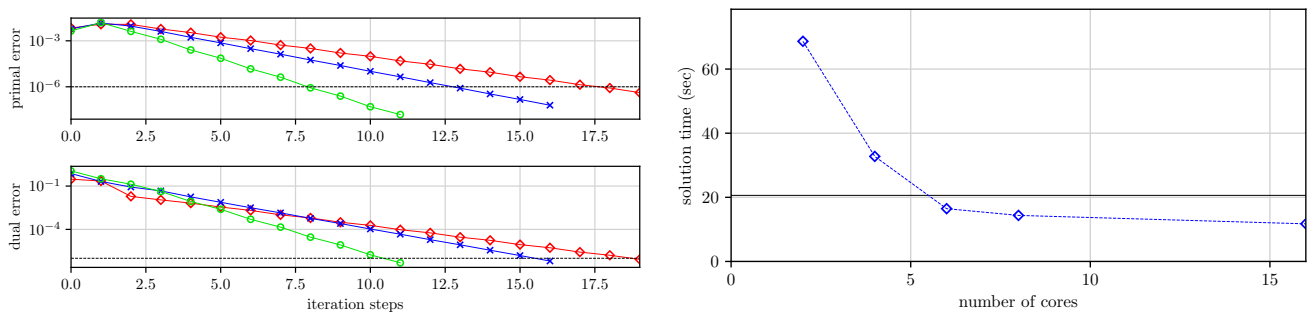


Fig. 4. Left: the evolution of the primal and dual residuals with  $\omega = 2$  (red and diamond),  $\omega = 4$  (blue and cross), and  $\omega = 8$  (green and circle) and with  $N = 4,800$ ,  $K = 8$ , and  $\rho_{\text{reg}} = 0.5$ . Right: a performance comparison of the centralized solver (black) and the proposed scheme with different number of cores (blue and diamond) and with  $N = 4,800$ ,  $\omega = 8$ , and  $\rho_{\text{reg}} = 0.5$ .

[21] J. Bailey, F. Horn, and R. Lin, "Cyclic operation of reaction systems: Effects of heat and mass transfer resistance," *AIChE Journal*, vol. 17, no. 4, pp. 818–825, 1971.

[22] I. Dunning, J. Huchette, and M. Lubin, "Jump: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.

[23] A. Wächter and L. T. Biegler, "On the implementation of an interior-

point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.