Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 49-7 (2016) 043–048

New Architectures for Hierarchical Predictive Control

Victor M. Zavala [*]

[*] Department of Chemical and Biological Engineering
University of Wisconsin-Madison
Madison, WI 53706 USA (e-mail: victor.zavala@wisc.edu).

Abstract: We analyze the structure of the Euler-Lagrange (EL) conditions of a long-horizon optimal control problem. The analysis reveals that the conditions can be solved by using block Gauss-Seidel (GS) schemes. We prove that such schemes can be implemented in the primal space by solving sequences of short-horizon optimal control problems. This analysis also reveals that a traditional receding-horizon (RH) scheme is equivalent to performing a single GS sweep. We have also found that we can use adjoint information from a coarse long-horizon problem to construct terminal penalties that correct the policies of the RH scheme. We observe that this scheme can be interpreted as a hierarchical controller in which a coarse high-level controller transfers long-horizon information to a low-level, short-horizon controller of fine resolution. The results open the door to a new family of hierarchical control architectures that can handle multiple time scales systematically.

Keywords: RH, hierarchical control, time scales, Euler-Lagrange

1. BASIC NOTATION AND SETTING

We start by providing basic notation and defining the technical problem. Relevant references are provided as we proceed with the discussion. We consider the following *long-horizon* optimal control problem:

$$\min_{z(\cdot),u(\cdot)} \int_0^T \varphi(z(\tau),u(\tau),w(\tau))d\tau \quad (1a)$$

$$\text{s.t.} \quad \dot{z}(\tau) = f(z(\tau),u(\tau),w(\tau)),\ \tau \in [0,T] \quad (1b)$$

$$z(0) = \bar{z}. \quad (1c)$$

Here, $z(\cdot)$, $u(\cdot)$, and $w(\cdot)$ are state, control, and disturbance trajectories, respectively. The cost and system mappings $\varphi(\cdot)$ and $f(\cdot)$ are assumed to be smooth.

We *lift* the long-horizon problem by partitioning the horizon T into n stages. This lifting approach was proposed by Bock and Plitt (1984) in the context of multiple-shooting. We define the sets $\mathcal{N} := \{0..n-1\}$ and $\mathcal{N}^- := \mathcal{N} \setminus \{n-1\}$; and we assume the stages to be of equal length $h := T/n$. The partitioning gives rise to the *lifted* problem,

$$\min_{z_k(\cdot),u_k(\cdot)} \sum_{k\in\mathcal{N}} \int_0^h \varphi(z_k(\tau),u_k(\tau),w_k(\tau))d\tau \quad (2a)$$

$$\text{s.t.} \ \dot{z}_k(\tau) = f(z_k(\tau),u_k(\tau),w_k(\tau)),\ k \in \mathcal{N},\ \tau \in [0,h] \quad (2b)$$

$$z_{k+1}(0) = z_k(h),\ k \in \mathcal{N}^- \quad (2c)$$

$$z_0(0) = \bar{z}. \quad (2d)$$

We will analyze the stage structure of the lifted optimal control problem. In doing so we will reduce the notation to a minimum, in such a way that it retains the essential features of the structure we are interested in highlighting. We first note that we do not consider inequality and path constraints and we eliminate dependencies of the mappings on the disturbances. These changes will not alter the stage structure of the lifted problem. We transcribe the lifted problem into a finite-dimensional nonlinear programming problem by applying an implicit Euler scheme with m inner stages of equal length $\delta := h/m$ (other discretization schemes can also be applied). We define the sets of inner discretization points $\mathcal{M} := \{0..m-1\}$. The discretized problem is,

$$\min_{z_{k,j},u_{k,j}} \sum_{k\in\mathcal{N}} \sum_{j\in\mathcal{M}} \varphi(z_{k,j+1},u_{k,j+1}) \quad (3a)$$

s.t.

$$(\nu_{k,j+1})\ z_{k,j+1} = z_{k,j} + \delta f(z_{k,j+1},u_{k,j+1}),\ k \in \mathcal{N}, j \in \mathcal{M} \quad (3b)$$

$$(\lambda_k) \quad z_{k,0} = z_{k-1,m},\ k \in \mathcal{N}. \quad (3c)$$

Here, $\nu_{k,j}$ are the dual variables of the inner dynamic equations (3b), and λ_k are the dual variables of the stage-transition equations (4b). The dual variables are scaled by the constant $1/\delta$. We use the dummy parameter $z_{-1,m} := \bar{z}$ to simplify notation. We denote the discretized long-horizon problem (4) as $\mathcal{P}$. We simplify notation further by eliminating the dynamic equations from the notation (3b). This, again, does not alter the stage structure. We obtain the compact problem,

$$\min_{z_{k,j},u_{k,j}} \sum_{k\in\mathcal{N}} \sum_{j\in\mathcal{M}} \varphi(z_{k,j+1},u_{k,j+1}) \quad (4a)$$

$$\text{s.t.} \quad (\lambda_k) \quad z_{k,0} = z_{k-1,m},\ k \in \mathcal{N}. \quad (4b)$$

A controller based on recursive solutions of $\mathcal{P}$ must capture disturbance signals that evolve over multiple time scales (e.g., noise, weather, prices) and must handle slow and fast components of the dynamical system (e.g., fast and slow chemical reactions, recycle systems). Despite advances in computational methods for optimal control, this might not be possible to do. This is is because the solution of $\mathcal{P}$ might require very fine discretization meshes and/or

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.07.214

expensive numerical integration procedures to capture dynamic effects at all time scales. Reviews on the topic are presented by Diehl et al. (2009) and Zavala and Biegler (2009). We also note that the presence of multiple time scales plays a role in the resolution and update frequency of the control. For instance, as noted by Findeisen et al. (2007), if disturbances are fast it is necessary to use a compatible control resolution.

The complexity of $\mathcal{P}$ is traditionally addressed by using a RH scheme which seeks to approximate the optimal long-horizon policy by solving sequences of fine-resolution short-horizon problems. In particular, one can solve the following short-horizon problems sequentially for $k = 0, ..., N-1$:

$$\min_{z_{k,j},u_{k,j}} \sum_{j\in\mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}) \tag{5a}$$

$$\text{s.t.} \quad (\lambda_k) \quad z_{k,0} = z_{k-1,m}. \tag{5b}$$

Here, the initial state $z_{k-1,m}$ is fixed and is obtained from the solution of the problem at $k-1$. We will show that this RH scheme is a block GS iteration applied to the solution of the Euler-Lagrange (EL) conditions of (4). This observation will help us derive hierarchical schemes to address the intractability of $\mathcal{P}$.

2. STRUCTURE OF EULER-LAGRANGE CONDITIONS

We group variables by stages by defining the vectors $\mathbf{z}_k := (z_{k,0}, ..., z_{k,m})$, $\mathbf{u}_k := (u_{k,1}, ..., u_{k,m})$, and $\nu_k := (\nu_{k,1}, ..., \nu_{k,m})$. We thus obtain the block form of $\mathcal{P}$,

$$\min_{\mathbf{u}_k} \sum_{k\in\mathcal{N}} \phi(\mathbf{z}_k, \mathbf{u}_k) \tag{6a}$$

$$\text{s.t.} \quad (\lambda_k) \quad \overline{\Pi}_k \mathbf{z}_k = \underline{\Pi}_k \, \mathbf{z}_{k-1}, \; k \in \mathcal{N}. \tag{6b}$$

The structure of the mapping $\phi(\cdot)$ is given by:

$$\phi(\mathbf{z}_k, \mathbf{u}_k) := \sum_{j\in\mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}). \tag{7}$$

The coefficient matrices $\overline{\Pi}_k$ and $\underline{\Pi}_k$ satisfy $\overline{\Pi}_k \mathbf{z}_k = z_{k,0}$ and $\underline{\Pi}_k \mathbf{z}_{k-1} = z_{k-1,m}$. We also define the fixed dummy vector $\mathbf{z}_{-1}$ satisfying $\underline{\Pi}_0 \, \mathbf{z}_{-1} = z_{-1,m} = \bar{z}$.

The Lagrange function of $\mathcal{P}$ is given by

$$\mathcal{L}(\mathbf{z}_k, \mathbf{u}_k, \lambda_k) := \sum_{k\in\mathcal{N}} \phi(\mathbf{z}_k, \mathbf{u}_k) - \lambda_k^T(\overline{\Pi}_k \mathbf{z}_k - \underline{\Pi}_k \mathbf{z}_{k-1}), \tag{8}$$

and its first-order optimality conditions are

$$0 = \nabla_z \phi_k - \overline{\Pi}_k^T \lambda_k + \underline{\Pi}_{k+1}^T \lambda_{k+1}, \; k \in \mathcal{N}^- \tag{9a}$$

$$0 = \nabla_z \phi_{n-1} - \overline{\Pi}_{n-1}^T \lambda_{n-1} \tag{9b}$$

$$0 = \nabla_u \phi_k, \; k \in \mathcal{N} \tag{9c}$$

$$0 = \overline{\Pi}_k \mathbf{z}_k - \underline{\Pi}_k \mathbf{z}_{k-1}, \; k \in \mathcal{N}. \tag{9d}$$

Here, $\nabla_z \phi_k := \nabla_{\mathbf{z}_k} \phi(\cdot)$ and $\nabla_u \phi_k := \nabla_{\mathbf{u}_k} \phi(\cdot)$. System (9) is the discrete-time version of the EL conditions of the lifted problem (2). Moreover, the dual variables λ_k can be tied together to form discrete-time profiles of the adjoint variables of the lifted problem. These properties are discussed in the book of Biegler (2010).

We note that the block component of the EL conditions corresponding to each stage $k \in \mathcal{N}^-$ is given by,

$$0 = \nabla_z \phi_k - \overline{\Pi}_k^T \lambda_k + \underline{\Pi}_{k+1}^T \lambda_{k+1} = 0 \tag{10a}$$

$$0 = \nabla_u \phi_k \tag{10b}$$

$$0 = \overline{\Pi}_k \mathbf{z}_k - \underline{\Pi}_k \mathbf{z}_{k-1} \tag{10c}$$

For fixed $\underline{\Pi}_k \mathbf{z}_{k-1} = z_{k-1,m}$ and λ_{k+1}, (10) are the first-order conditions of the primal stage problem:

$$\min_{\mathbf{z}_k,\mathbf{u}_k} \quad \phi(\mathbf{z}_k, \mathbf{u}_k) + (\lambda_{k+1})^T \underline{\Pi}_{k+1} \mathbf{z}_k \tag{11a}$$

$$\text{s.t.} \quad (\lambda_k) \quad \overline{\Pi}_k \mathbf{z}_k = \underline{\Pi}_k \, \mathbf{z}_{k-1}. \tag{11b}$$

For the last stage $k = n-1$ we have the block component of the EL conditions:

$$0 = \nabla_z \phi_{n-1} - \overline{\Pi}_{n-1}^T \lambda_{n-1} = 0 \tag{12a}$$

$$0 = \nabla_u \phi_{n-1} \tag{12b}$$

$$0 = \overline{\Pi}_{n-1} \mathbf{z}_{n-1} - \underline{\Pi}_{n-1} \mathbf{z}_{n-2}. \tag{12c}$$

For fixed $\underline{\Pi}_{n-1} \mathbf{z}_{n-2} = z_{n-2,m}$ these are the first-order conditions of the primal stage problem,

$$\min_{\mathbf{z}_{n-1},\mathbf{u}_{n-1}} \quad \phi(\mathbf{z}_{n-1}, \mathbf{u}_{n-1}) \tag{13a}$$

$$\text{s.t.} \quad (\lambda_{n-1}) \quad \overline{\Pi}_{n-1} \mathbf{z}_{n-1} = \underline{\Pi}_{n-1} \, \mathbf{z}_{n-2}. \tag{13b}$$

From the structure of (10) and (11) we can see that coupling between neighboring stages $k-1$, k, and $k+1$ is introduced through the states $z_{k-1,m}$ and adjoints λ_{k+1}.

3. BLOCK GS SCHEMES

Our key observation is that we make is that we can solve the EL conditions (9) of the long-horizon problem by using *block* GS schemes. Assume that the adjoints λ_k are fixed to $\lambda_k^\ell = 0$ for all $k \in \mathcal{N}$. At stage $k = 0$ and with fixed $z_{-1}^\ell = \bar{z}$ we solve the short-horizon problem:

$$\min_{z_{k,j},u_{k,j}} \sum_{j\in\mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}) + \delta(\lambda_{k+1}^\ell)^T z_{k,m} \tag{14a}$$

$$\text{s.t.} \quad (\lambda_k) \quad z_{k,0} = z_{k-1,m}^\ell. \tag{14b}$$

We refer to this problem as $\mathcal{P}_k$ and introduce the notation

$$(z_{k,m}^{\ell+1}, \lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k-1,m}^\ell, \lambda_{k+1}^\ell) \tag{15}$$

to indicate the inputs and outputs of problem $\mathcal{P}_k$. The primal-dual solution of $\mathcal{P}_k$ solves block k of the EL conditions (10) for fixed initial state $\underline{\Pi}_k \mathbf{z}_{k-1} = z_{k-1,m}^\ell$ and adjoint $\lambda_{k+1} = \lambda_{k+1}^\ell$. Note also that $\mathcal{P}_k$ is equivalent to the stage problem (11).

From the solution of $\mathcal{P}_k$ we obtain the terminal state $z_{k,m}^{\ell+1}$ and we use this as initial state for $\mathcal{P}_{k+1}$ to compute $(z_{k+1,n}^{\ell+1}, \lambda_{k+1}^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k,m}^{\ell+1}, \lambda_{k+2}^\ell)$. We continue the recursion until reaching the last stage, $k = n-1$. At this stage we solve problem $\mathcal{P}_{n-1}$:

$$\min_{z_{n-1,j},u_{n-1,j}} \sum_{j\in\mathcal{M}} \varphi(z_{n-1,j+1}, u_{n-1,j+1}) \tag{16a}$$

$$\text{s.t.} \quad (\lambda_{n-1}) \quad z_{n-1,0} = z_{n-2,m}^\ell. \tag{16b}$$

With this we compute $(z_{n-1,m}^{\ell+1}, \lambda_{n-1}^{\ell+1}) \leftarrow \mathcal{P}_{n-1}(z_{n-2,m}^\ell, 0)$. The primal-dual solution of $\mathcal{P}_{n-1}$ solves the optimality system (12) for fixed initial state $\underline{\Pi}_{n-1} \mathbf{z}_{n-2} = z_{n-2,m}^\ell$ obtained from the solution of $\mathcal{P}_{n-2}$. Moreover, $\mathcal{P}_{n-1}$ is equivalent to (13).

After solving $\mathcal{P}_{n-1}$ we have updated all the state (primal) $z_k^{\ell+1}$ and adjoint $\lambda_k^{\ell+1}$ variables. In Figure 1 we can see

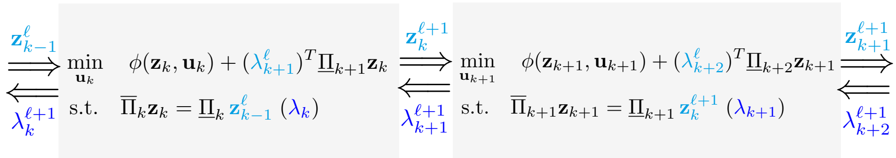

Fig. 1. Sketch of primal and dual updates in GS scheme.

that the updates of the states march forward in time while the updates march backward in time. After this sweep, we return to the first stage $k=0$ and repeat the recursion to obtain $z_k^{\ell+2},\lambda_k^{\ell+2}$. We repeat this procedure n_{GS} times. We summarize the GS scheme below:

GS Scheme

I) GIVEN $\bar{z}$, set counter $\ell \leftarrow 0$, set $z_{-1,m}^{\ell} \leftarrow \bar{z}$, and set $\lambda_k^{\ell} \leftarrow 0$ for $k=0,...,n-1$. FOR $\ell=0,...,n_{GS}$ DO:
II) FOR $k=0,...,n-2$ SOLVE $(z_{k,m}^{\ell+1},\lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k-1,m}^{\ell+1},\lambda_{k+1}^{\ell})$.
III) FOR $k=n-1$ SOLVE $(z_{n-1,m}^{\ell+1},\lambda_{n-1}^{\ell+1}) \leftarrow \mathcal{P}_{n-1}(z_{n-2,m}^{\ell},0)$.
IV) SET $\ell \leftarrow \ell+1$ and RETURN TO Step II).

We note that the *GS scheme can be implemented by using off-the-shelf optimization tools*. All that is needed to solve the EL conditions (9) is the ability to solve the NLPs $\mathcal{P}_k$. The structure of the block GS scheme also reveals that a *RH scheme is equivalent to performing a single GS iteration* with adjoints $\lambda_k^{\ell}=0,\ k\in\mathcal{N}$.

The adjoints λ_{k+1}^{ℓ} encode important global information of the future horizon beyond the short-horizon h of $\mathcal{P}_k$. In particular, we note that the adjoints can be interpreted as *terminal costs*. The terminal term $(\lambda_{k+1}^{\ell})^T z_{k,m}$ in $\mathcal{P}_k$ is a first-order approximation of the so-called *cost-to-go*. To see this, consider a problem with two stages k and $k+1$:

$$\min_{\mathbf{z}_k,\mathbf{u}_k,\mathbf{z}_{k+1},\mathbf{u}_{k+1}} \quad \phi(\mathbf{z}_k,\mathbf{u}_k)+\phi(\mathbf{z}_{k+1},\mathbf{u}_{k+1}) \tag{17a}$$

$$\text{s.t.} \quad \overline{\Pi}_k\mathbf{z}_k=\underline{\Pi}_k\,\mathbf{z}_{k-1}^{\ell} \quad (\lambda_k) \tag{17b}$$

$$\overline{\Pi}_{k+1}\mathbf{z}_{k+1}=\underline{\Pi}_{k+1}\,\mathbf{z}_k \quad (\lambda_{k+1}) \tag{17c}$$

This problem can be written as:

$$\min_{\mathbf{z}_k,\mathbf{u}_k} \ \phi(\mathbf{z}_k,\mathbf{u}_k)+\mathcal{Q}(\underline{\Pi}_{k+1}\mathbf{z}_k) \tag{18a}$$

$$\text{s.t.} \quad \overline{\Pi}_k\mathbf{z}_k=\underline{\Pi}_k\,\mathbf{z}_{k-1}^{\ell} \quad (\lambda_k) \tag{18b}$$

where,

$$\mathcal{Q}(\underline{\Pi}_{k+1}\,\mathbf{z}_k):=\min_{\mathbf{z}_{k+1},\mathbf{u}_{k+1}} \quad \phi(\mathbf{z}_{k+1},\mathbf{u}_{k+1}) \tag{19a}$$

$$\text{s.t.} \quad \overline{\Pi}_{k+1}\mathbf{z}_{k+1}=\underline{\Pi}_{k+1}\,\mathbf{z}_k\ (\lambda_{k+1}). \tag{19b}$$

Here $\mathcal{Q}(\cdot)$ is the cost-to-go function. If we linearize the cost-to-go at the current state guess $\mathbf{z}_k^{\ell}$ we obtain,

$$\min_{\mathbf{z}_k,\mathbf{u}_k} \ \phi(\mathbf{z}_k,\mathbf{u}_k) + \mathcal{Q}(\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell})+\partial\mathcal{Q}(\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell})^T(\underline{\Pi}_{k+1}\mathbf{z}_k-\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell}) \tag{20a}$$

$$\text{s.t.} \quad \overline{\Pi}_k\mathbf{z}_k=\underline{\Pi}_k\,\mathbf{z}_{k-1}^{\ell} \quad (\lambda_k) \tag{20b}$$

where the term $\partial\mathcal{Q}(\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell})^T\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell}$ is fixed and therefore irrelevant. From duality we know that the gradient of the cost-to-go $\partial\mathcal{Q}(\underline{\Pi}_{k+1}\mathbf{z}_k^{\ell})$ is precisely the adjoint λ_{k+1}^{ℓ} of the state transition constraint (19b). In other words, the adjoint is the sensitivity of the cost-to-go with respect to the initial state $\underline{\Pi}_{k+1}\,\mathbf{z}_k$. Consequently, problem (20) is equivalent to,

$$\min_{\mathbf{z}_k,\mathbf{u}_k} \ \phi(\mathbf{z}_k,\mathbf{u}_k)+(\lambda_{k+1}^{\ell})^T\underline{\Pi}_{k+1}\mathbf{z}_k \tag{21a}$$

$$\text{s.t.} \quad \overline{\Pi}_k\mathbf{z}_k=\underline{\Pi}_k\,\mathbf{z}_{k-1}^{\ell} \quad (\lambda_k). \tag{21b}$$

This is precisely the k stage subproblem $\mathcal{P}_k$ solved in the GS scheme. We also note that the adjoints are exact approximations of the cost-to-go at an optimal solution of $\mathcal{P}$. The cost-to-go information is *neglected* by the RH scheme and this can lead to a poor approximation of the long-horizon solution. A key insight that we gain from our analysis and from the cost-to-go interpretation is that we *can correct the RH scheme* to better approximate the long-horizon solution if we are capable of obtaining estimates of the adjoints λ_k^{ℓ}. Moreover, in the ideal case where such adjoint estimates are optimal for $\mathcal{P}$, the corrected RH scheme will deliver optimal long-horizon profiles for $\mathcal{P}$.

4. COARSENING-BASED CORRECTION

We can obtain estimates of the adjoints λ_k^{ℓ} by *performing multiple GS* iterations. This can be seen as a *self-correcting* RH scheme. GS schemes provide the computational advantage that they only need to solve only short-horizon problems. GS schemes, however, are well known for exhibiting slow convergence or no convergence at all. We address this issue by using adjoint estimates λ_k^{ℓ} obtained from the solution of a *coarsened* long-horizon problem. To perform coarsening, we consider a coarse grid with m_c elements and $m_c \ll m$ (such that the resulting coarse problem is tractable). We define the coarse set as $\mathcal{M}_c := \{0..m_c-1\}$ and the corresponding coarsened long-horizon problem $\mathcal{P}_c$. We use the notation $(\lambda_0^{\ell},...,\lambda_{n-1}^{\ell}) \leftarrow \mathcal{P}_c(\bar{z})$ to indicate the inputs and outputs of $\mathcal{P}_c$. We consider the scheme:

Corrected GS Scheme

I) GIVEN $\bar{z}$, set counter $\ell \leftarrow 0$, set $z_{-1,m}^{\ell} \leftarrow \bar{z}$.
II) SOLVE $(\lambda_0^{\ell},...,\lambda_{n-1}^{\ell}) \leftarrow \mathcal{P}_c(\bar{z})$ and FOR $\ell = 0,...,n_{GS}$ DO:
III) FOR $k=0,...,n-2$ SOLVE $(z_{k,m}^{\ell+1},\lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k-1,m}^{\ell+1},\lambda_{k+1}^{\ell})$.
IV) FOR $k=n-1$ SOLVE $(z_{n-1,m}^{\ell+1},\lambda_{n-1}^{\ell+1}) \leftarrow \mathcal{P}_{n-1}(z_{n-2,m}^{\ell},0)$.
V) SET $\ell \leftarrow \ell+1$ and RETURN TO Step III).

In this scheme, $\mathcal{P}_c$ transfers global (long-horizon) information of problem $\mathcal{P}$ to the local (short-horizon) problems $\mathcal{P}_k$. This is depicted in Figure 2. This approach can be seen as a *hierarchical control scheme* in which a coarse-grained high-level controller supervises a fine-grained low-level controller. In other words, the coarse long-horizon problem provides terminal costs (i.e., a cost-to-go) to the RH controller so that this better approximates the solution of the long-horizon problem. Dual information is

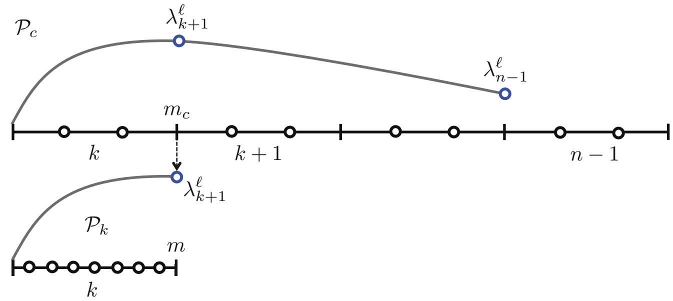

Fig. 2. Transfer of adjoint information from coarse long-horizon problem to fine short-horizon problems.

transferred in this hierarchical controller; *as opposed to traditional hierarchical schemes that transfer state information*. This is key because dual information provided by the coarse controller can be refined by the low-level controller (through GS updates) and thus it is possible to achieve *primal-dual optimality*.

In the comprehensive review on hierarchical predictive control of Scattolini (2009), the author notices that systematic design methods for hierarchical control are still lacking. More specifically, no hierarchical schemes have been proposed that systematically aggregate and refine trajectories at multiple time scales. In addition, existing schemes have been tailored to achieve feasibility but *do not have optimality guarantees*. Scattolini also argues that it is important to consider multirate methods (computing controls at different time resolutions). As can be seen, the hierarchical control structure created in this work addresses both of these issues. It is also worth remarking that several hierarchical schemes for MPC have been proposed that decouple time scales using singular perturbations. For more information on such approaches, the reader is referred to the work of Ellis et al. (2013); Chen et al. (2012); Scattolini (2009). These approaches enable handling of slow and fast modes but do not address multiple time scales.

Our analysis borrows concepts of multigrid control. We now describe a general multigrid framework to highlight the many connections that exist with hierarchical control. Multigrid is a *scientific computing* paradigm widely used for the simulation and control of systems described by PDEs Borzı (2003). The idea behind multigrid is to discretize the PDE *space-time domain* in a sequence of grids of increasing resolution. At the highest level is the *coarsest* grid that is computationally tractable. The key observation is that the coarse level can capture slow moving fronts occurring at low frequencies but will miss the fast moving fronts occurring at higher frequencies. The coarse state, control, and adjoint (dual or co-state) fields are passed to the second level of the hierarchy. At this level, the coarse fields are refined using a smoothing scheme (typically a GS scheme). The GS scheme exchanges information among its nearest neighbors. By doing this, the GS scheme can eliminate fast local moving fronts quickly (but not necessarily the slow global ones). Consequently, by combining the coarse and fine levels one seeks to target the entire frequency spectrum. As can be seen, many of the problems arising in the simulation and control of PDEs also appear in hierarchical control. We have addressed several technical issues that enabled us to transfer ideas between these two domains. These include:

- We established connections between the structure of the EL conditions and the solution of stage-by-stage optimal control problems in the primal space. This enables us to use GS schemes in a general setting and establishes a connection between RH control and GS.
- We use a primal representation of stage problems $\mathcal{P}_k$ corresponding to the stage block of the EL conditions. With this we demonstrate that coarsening schemes can be used to obtain adjoint information that is used as terminal penalties of short-horizon stage problems.
- We demonstrate that GS and coarsening schemes can be implemented using off-the-shelf optimization tools, as opposed to intrusive linear algebra manipulations used in PDE studies.
- We demonstrate that only dual information needs to be communicated between hierarchical levels, as opposed to only state and control information communicated in multigrid PDE schemes.
- The lifting approach avoids the need to perform interpolation of state and adjoint profiles in order to move from a coarse grid to a fine grid (as is typically done in PDEs). In the proposed approach, all that is needed from the coarse problem are the adjoints at the stage transition points.

5. NUMERICAL STUDIES

We discuss the numerical concepts using a CSTR example and we then demonstrate how to use hierarchical control in a more complicated microgrid control example.

5.1 CSTR

We consider the nonlinear Hicks-Ray CSTR reactor problem reported in Zavala and Anitescu (2010). The system states are the concentration of reactant $c(\cdot)$ and the temperature of reacting mixture $t(\cdot)$. The control is the cooling water flow $u(\cdot)$. After lifting, we denote the adjoint associated with the concentration as λ_c and the adjoint associated with temperature as λ_t. We partition the full time horizon in $n = 10$ stages and discretize each stage using an implicit Euler scheme with $m = 10$ grid points. In Figure 3 we present the control profiles obtained with the GS Scheme. As can be seen, the profiles obtained with a standard RH scheme (first iteration of GS scheme) severely deviate from the optimal ones. GS approximates the long-horizon solution fairly well after three iterations.

In Figure 4 we present the optimal and approximate adjoint profiles obtained from coarsening. The coarse profiles are obtained by discretizing the stage using $m_c = 2$ points (compared to the $m = 10$ points used for the optimal profile). This is a coarsening factor of five. As can be seen, the adjoint profiles exhibit errors but the overall structure is preserved. We use the coarse adjoint profiles to correct the GS scheme (provide terminal penalties for the short-horizon stage problems). In Figure 5 we present the convergence of GS with and without correction. As can be seen, coarse correction delivers close-to-optimal performance optimal performance after a single GS iteration.

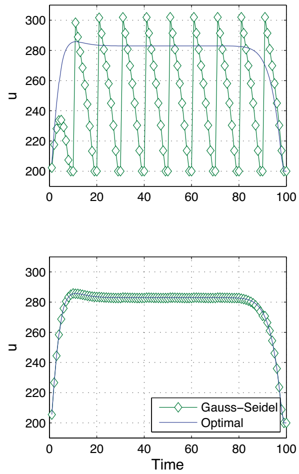

Fig. 3. Scheme convergence for control. First iteration (top) and third iteration (bottom).

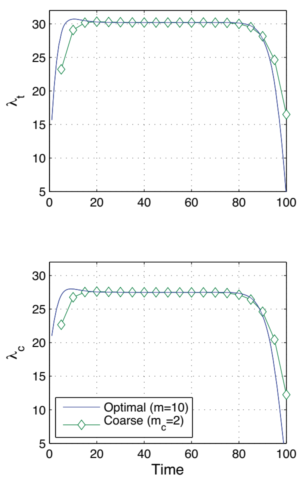

Fig. 4. Optimal and coarsened adjoint profiles for temperature (top) and concentration (bottom).

6. MICROGRID

We now demonstrate how the hierarchical controller can manage disturbances with multiple time scales. We consider the *microgrid system* shown in Figure 6. The system has two generators and a battery storage system. The objective to find the optimal operating policy for the generators and storage that minimizes cost over a long horizon spanning 4 days. We assume that the electrical load (disturbance) of this system has three frequency components (load has periods of 24 hr, 12 hr, and 1 hr). These are shown in the right panel of Figure 7. In the left panels of Figure 7 we show the coarsened load profile and the error induced by applying such coarsening. As can be seen, coarsening misses the high frequency.

In the left column of Figure 8 we present, from top to bottom, the storage policies for short-horizon RH control (GS without adjoint information), GS scheme after two iterations, coarse control, and hierarchical control. We split the horizon in 4 stages of one day and coarsen the horizon by a factor of four. In the right column we present the suboptimality errors for the corresponding storage policies with respect to the long-horizon problem $\mathcal{P}$. From the first

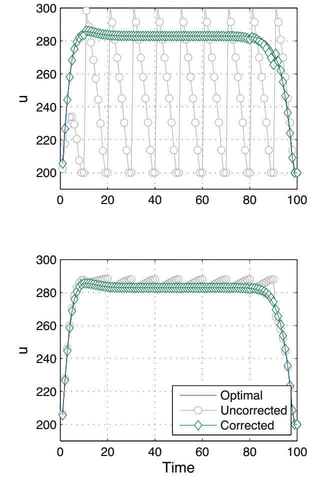

Fig. 5. Control policies for GS scheme with and without coarsening correction. First iteration (top) and second iteration (bottom).

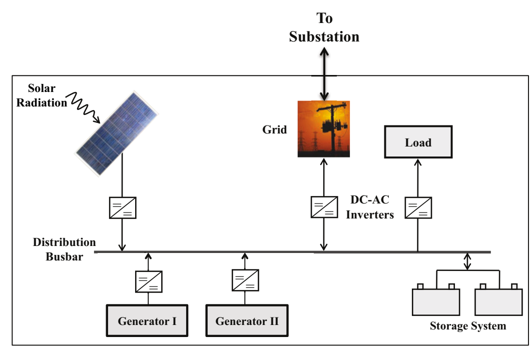

Fig. 6. Microgrid control example.

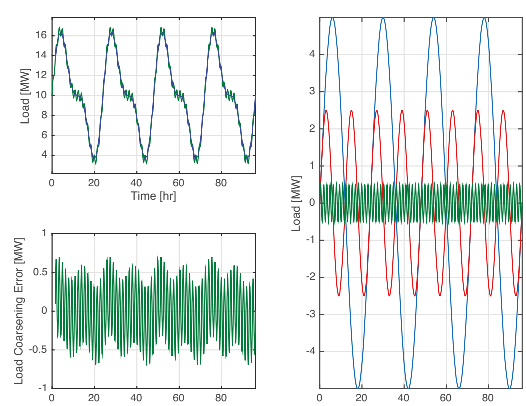

Fig. 7. Real and coarsened loads (top left), load coarsening error (bottom left), and frequency components of load (right).

row we see that the *RH policy is far from optimal*. This is because the controller marches in stages of one day and thus ignores long horizon load trends. As a result, the RH controller depletes the current inventory in trying to minimize short-term generation cost. In the second row we see that a GS scheme can better approximate the optimal policy after the second iteration because it can estimate the adjoints that are used to capture long horizon information. The error in the storage policy is large in

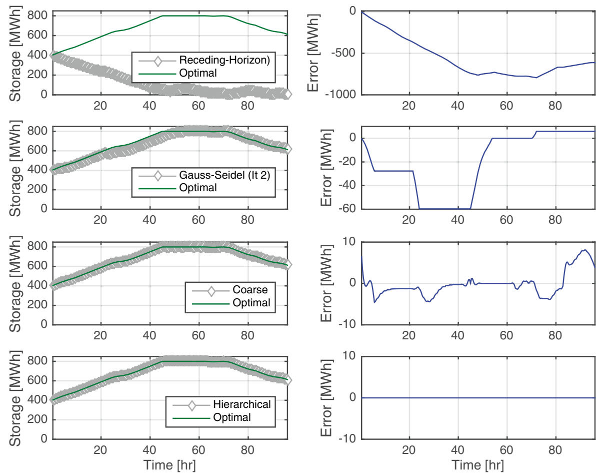

Fig. 8. From top to bottom: Storage policies and errors with short-horizon RH scheme, GS scheme (after 2 iterations), coarse control, and hierarchical control.

magnitude but it is *flattened out*, indicating that the GS scheme eliminates the high frequencies of the load but not the low frequencies. In the third row we see that a pure *coarse controller* has close-to-optimal performance but the storage error shows that it *misses the high frequencies* of the load. In the fourth row we see that hierarchical control obtains optimal performance because it can deal with high and low frequencies. In Figure 9 we present storage, generation, and adjoint policies. We again note that RH policies are far from optimal and that the GS scheme can recover them after a few iterations.

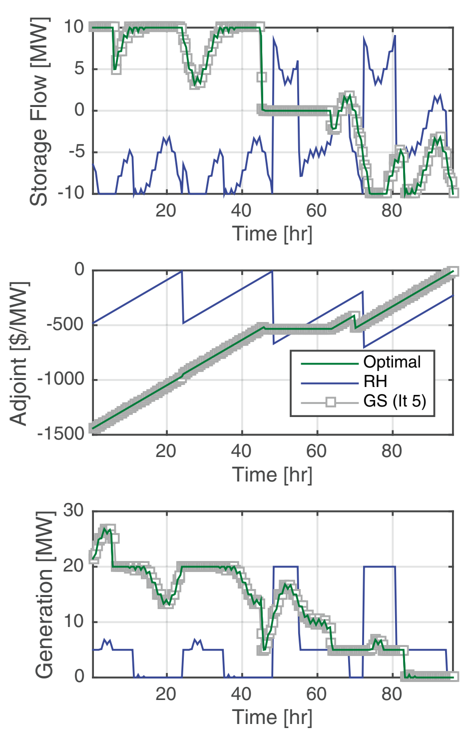

Fig. 9. From top to bottom: storage, adjoint, and generation profiles.

7. CONCLUSIONS AND FUTURE WORK

We presented an analysis of the Euler-Lagrange conditions for a lifted optimal control problem. This enabled us to derive block GS schemes capable of solving intractable long-horizon problems by solving sequences of tractable short-horizon problems. Our analysis revealed that a RH scheme is equivalent to performing a GS sweep of the Euler-Lagrange conditions. We have also used our analysis to derive strategies to correct adjoint profiles by using coarsening. This approach can interpreted as a hierarchical control structure in which a coarse high-level controller transfers long-horizon information to a low-level, short-horizon controller of fine resolution. Our results open the door to the design of new hierarchical control applications. As part of future work, we would like to gain additional insight on conditions guaranteeing convergence of GS schemes and we will design multi-level hierarchies .

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, under an Early Career Award. The author thanks Mihai Anitescu for pointing out existing work on multigrid control.

REFERENCES

Biegler, L.T. (2010). *Nonlinear programming: Concepts, algorithms, and applications to chemical processes*. SIAM.

Bock, H.G. and Plitt, K.J. (1984). A multiple shooting algorithm for direct solution of optimal control problems. *9th IFAC World Congress, Budapest*.

Borzı, A. (2003). Multigrid methods for parabolic distributed optimal control problems. *Journal of Computational and Applied Mathematics*, 157(2), 365–382.

Chen, X., Heidarinejad, M., Liu, J., and Christofides, P.D. (2012). Composite fast-slow mpc design for nonlinear singularly perturbed systems. *AIChE Journal*, 58(6), 1802–1811.

Diehl, M., Ferreau, H.J., and Haverbeke, N. (2009). Efficient numerical methods for nonlinear MPC and moving horizon estimation. *In Nonlinear Model Predictive Control*, 391–417.

Ellis, M., Heidarinejad, M., and Christofides, P.D. (2013). Economic model predictive control of nonlinear singularly perturbed systems. *Journal of Process Control*, 23(5), 743–754.

Findeisen, R., Raff, T., and Allgöwer, F. (2007). Sampled-data nonlinear model predictive control for constrained continuous time systems. In *Advanced Strategies in Control Systems with Input and Output Constraints*, 207–235. Springer.

Scattolini, R. (2009). Architectures for distributed and hierarchical model predictive control–A review. *Journal of Process Control*, 19(5), 723–731.

Zavala, V.M. and Biegler, L.T. (2009). Nonlinear programming strategies for state estimation and model predictive control. *In Nonlinear Model Predictive Control*, 419–432.

Zavala, V.M. and Anitescu, M. (2010). Real-time nonlinear optimization as a generalized equation. *SIAM Journal on Control and Optimization*, 48(8), 5444–5467.