

Selective Input Adaptation in Parametric Optimal Control Problems involving Terminal Constraints

Saurabh Deshpande, Dominique Bonvin, and Benoît Chachuat

Abstract—This paper is concerned with input adaptation in dynamic processes in order to guarantee feasible and optimal operation despite the presence of uncertainty. For optimal control problems having terminal constraints, two sets of directions can be distinguished in the input function space: the so-called sensitivity-seeking directions, along which a small input variation does not affect the terminal constraints, and the complementary constraint-seeking directions, along which a variation does affect the terminal constraints. Two selective input adaptation scenarios are thus possible, namely, adaptation along each set of input directions. This paper proves the important result that the cost variation due to the adaptation along the sensitivity-seeking directions is typically smaller than that due to the adaptation along the constraint-seeking directions.

Index Terms—Parametric optimal control, terminal constraints, sensitivity-seeking directions, constraint-seeking directions, Fredholm integral equations, selective input adaptation.

I. INTRODUCTION

Processes that are either inherently transient or operated in an unsteady-state manner are abundant in the resource industries. Examples in chemical engineering include batch and semi-batch processes that are characterized by the absence of a steady state. In the energy sector, transient systems are also expected to play a key role in future years, as many alternative technologies rely on discontinuous operation.

We will consider the problem of optimal control of transient processes for which the uncertainties in the process model are represented in the form of parametric variations. The optimal input profiles are typically calculated off-line, before the process starts, and are then applied to the process in an open-loop manner. Naturally, when some parameters deviate from their nominal values during the process operation, a change in optimal inputs is required to maintain optimality and meet operational constraints. Adapting *all* parts of the optimal input profiles to compensate for the effect of parametric variations is rarely possible in practice, nor is it expedient from a performance viewpoint. Partial or selective input adaptation scenarios that result in acceptable performance loss compared to optimal operation of the perturbed process are therefore worth considering.

For problems comprising *path constraints*, the possibility of splitting the input space, at each time instant, into the

so-called (pointwise) *sensitivity- and constraint-seeking* directions has been demonstrated in [1]. It follows that, for example, a small input variation along the former set of directions at a given time does not change the active path constraint values at that time. The definition of these input directions is thus tied to the variation in constraints due to instantaneous input changes. On the other hand, in the presence of *terminal constraints*, one must anticipate the effect of input adaptation at every intermediate time on the terminal constraint values. Accordingly, defining constraint- and sensitivity-seeking directions in the pointwise sense for terminal-constrained problems is not possible, thereby making the design of selective input adaptation schemes more challenging.

The purpose of this paper is to address this important challenge. A new definition of sensitivity- and constraint-seeking directions is developed, which considers the changes in terminal constraint values resulting from *all* input variations along the optimization horizon. In this approach, a sensitivity-seeking direction turns out to be the solution of a particular linear Fredholm integral equation of the first kind. In other words, the sensitivity- and constraint-seeking directions are now *directions* in the input function space $\mathcal{C}[t_0, t_f]^{n_u}$ as opposed to directions in the finite-dimensional space \mathbb{R}^{n_u} at each time instant as in [1]. It is then proved that the cost variation (with respect to no adaptation of the nominal optimal inputs) achieved by making the adaptation along the sensitivity-seeking directions is smaller than when making the adaptation along the constraint-seeking directions.

The outline of the paper is as follows. The mathematical formulation of the parametric optimal control problem involving terminal constraints is given in Section II, along with a summary of the necessary conditions of optimality (NCOs). In Section III, the sensitivity- and constraint-seeking directions are defined, and the concept of selective input adaptation along each set of directions is introduced. Section IV presents a quantitative comparison of the cost variation due to either of these two selective input adaptation scenarios. A numerical procedure to compute a specific set of constraint- and sensitivity-seeking directions is discussed in Section V and is illustrated by means of an example in Section VI. Finally, Section VII summarizes the results and identifies future research directions.

II. PARAMETRIC OPTIMAL CONTROL PROBLEM

The following parametric optimal control problem in the parameters θ , subject to the terminal inequality constraints

Saurabh Deshpande and Dominique Bonvin are with the Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. {saurabh.deshpande, dominique.bonvin}@epfl.ch
Benoît Chachuat is with the Department of Chemical Engineering, McMaster University, Hamilton, ON L8S 4L7, Canada. benoit@mcmaster.ca

$\mathbf{T} \leq \mathbf{0}$, with given initial time t_0 and terminal time t_f , is considered (OC(θ)): ¹

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}); \quad \mathbf{x}(t_0) = \mathbf{h}(\boldsymbol{\theta}), \quad (1)$$

$$T_i(t_f, \mathbf{x}(t_f), \boldsymbol{\theta}) \leq 0, \quad i \in \mathcal{I}_{n_T}, \quad (2)$$

$$\min_{\mathbf{u}} J = \psi(t_f, \mathbf{x}(t_f), \boldsymbol{\theta}) + \int_{t_0}^{t_f} \phi(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}) dt, \quad (3)$$

where $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ and $\mathbf{x}(t) \in \mathbb{R}^{n_x}$.

The functions \mathbf{f} , \mathbf{T} , ψ and ϕ in (OC(θ)) are assumed to be continuously differentiable with respect to all their arguments.

Let the nominal values of the system parameters be $\boldsymbol{\theta}_0$, and let $(\mathbf{u}^*(t), \mathbf{x}^*(t))$ be an optimal pair for the problem OC($\boldsymbol{\theta}_0$). We assume the following constraint qualification to hold ([2]): $\text{rank}\{\mathbf{T}_x^a(t_f, \mathbf{x}^*(t_f), \boldsymbol{\theta}_0)\} = n_{\mathbf{T}^a}$, where $n_{\mathbf{T}^a}$ denotes the number of active terminal constraints.² Introducing the Hamiltonian function \mathcal{H} ,

$$\mathcal{H}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\theta}) := \phi(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) + \boldsymbol{\lambda}^T \mathbf{f}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\theta}),$$

and assuming that the problem OC($\boldsymbol{\theta}_0$) is not abnormal, the so-called first-order necessary condition of optimality must hold almost everywhere (a.e.) in $[t_0, t_f]$ [2]:

$$\mathbf{0} = \mathcal{H}_{\mathbf{u}}(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), \boldsymbol{\theta}_0), \quad (4)$$

$$\dot{\boldsymbol{\lambda}}^*(t) = -\mathcal{H}_{\mathbf{x}}(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), \boldsymbol{\theta}_0) \quad (5)$$

$$\boldsymbol{\lambda}^*(t_f) = \psi_{\mathbf{x}}(t_f, \mathbf{x}^*(t_f), \boldsymbol{\theta}_0) + \mathbf{T}_{\mathbf{x}}(t_f, \mathbf{x}^*(t_f), \boldsymbol{\theta}_0)^T \boldsymbol{\rho}^*$$

$$0 = \rho_i^* T_i(t_f, \mathbf{x}^*(t_f), \boldsymbol{\theta}_0), \quad \forall i \in \mathcal{I}_{\mathbf{T}}$$

$$0 \leq \rho_i^*, \quad \forall i \in \mathcal{I}_{\mathbf{T}}.$$

for some multiplier functions $\boldsymbol{\lambda}^*(t) \in \mathbb{R}^{n_x}$, and multipliers $\boldsymbol{\rho}^* \in \mathbb{R}^{n_T}$.

For the analysis that will follow, we will make two more assumptions:

- The following *strict complementarity condition* holds: The multipliers ρ_i^* corresponding to the *active* terminal constraints are strictly nonzero; the vector of these multipliers will henceforth be denoted by $\boldsymbol{\rho}_a$.
- The Hamiltonian function is *regular*, which implies that the optimal inputs \mathbf{u}^* are continuous in $[t_0, t_f]$.

During process operation, the value of the system parameters can deviate from their nominal values $\boldsymbol{\theta}_0$. To compensate the effect of such variations, it becomes necessary to adapt the input profiles in such a way that they satisfy the optimality conditions for the perturbed problem.

III. CONSTRAINT- AND SENSITIVITY-SEEKING DIRECTIONS

In this section, the sensitivity- and constraint-seeking directions in input space are characterized by considering small variations of a specific type in the optimal inputs around their nominal optimal values \mathbf{u}^* . The characterization of the sensitivity- and constraint-seeking directions for the

problem under consideration will be based on the variation in the values of the active terminal constraints due to the aforementioned small input variations.

Consider a small variation around the nominal optimal inputs of the form

$$\tilde{\mathbf{u}}(t) = \mathbf{u}^*(t) + \eta \boldsymbol{\xi}^u(t), \quad |\eta| \ll 1. \quad (6)$$

Henceforth, we will say that such an input variation is *along the direction* $\boldsymbol{\xi}^u$. Let the resulting perturbed states be denoted by $\tilde{\mathbf{x}}(t)$. Thus, the pair $(\tilde{\mathbf{x}}(t), \tilde{\mathbf{u}}(t))$ satisfies (1) for $\boldsymbol{\theta}_0$, and we have ³

$$\dot{\tilde{\mathbf{x}}}(t) - \dot{\mathbf{x}}^*(t) = \mathbf{f}(t, \tilde{\mathbf{x}}(t), \tilde{\mathbf{u}}(t), \boldsymbol{\theta}_0) - \mathbf{f}(t).$$

Because of the continuous differentiability of \mathbf{f} with respect to the inputs and states at $(\mathbf{u}^*(t), \mathbf{x}^*(t))$, we can consider the Taylor expansion of \mathbf{f} around $(\mathbf{u}^*(t), \mathbf{x}^*(t))$ to obtain:

$$\begin{aligned} \frac{d}{dt} \{\tilde{\mathbf{x}}(t) - \mathbf{x}^*(t)\} \\ = \mathbf{f}_{\mathbf{x}}[t] \{\tilde{\mathbf{x}}(t) - \mathbf{x}^*(t)\} + \eta \mathbf{f}_{\mathbf{u}}[t] \boldsymbol{\xi}^u(t) + O(\eta^2). \end{aligned}$$

This yields the following first-order approximation of $\tilde{\mathbf{x}}(t; \eta)$:

$$\tilde{\mathbf{x}}(t; \eta) = \mathbf{x}^*(t) + \eta \boldsymbol{\xi}^x(t) + O(\eta^2), \quad (7)$$

where $\boldsymbol{\xi}^x(t)$ is the solution of

$$\begin{aligned} \dot{\boldsymbol{\xi}}^x(t) &= \mathbf{f}_{\mathbf{x}}[t] \boldsymbol{\xi}^x(t) + \mathbf{f}_{\mathbf{u}}[t] \boldsymbol{\xi}^u(t), \quad \forall t \in [t_0, t_f], \\ \boldsymbol{\xi}^x(t_0) &= \mathbf{0}. \end{aligned}$$

In particular, the unique solution to the above linear system can be written in the form [3]

$$\boldsymbol{\xi}^x(t) = \int_{t_0}^t \boldsymbol{\Phi}^{\mathbf{f}_x}(t, s) \mathbf{f}_{\mathbf{u}}[s] \boldsymbol{\xi}^u(s) ds, \quad \forall t \in [t_0, t_f], \quad (8)$$

where $\boldsymbol{\Phi}^{\mathbf{A}}(t, s)$ stands for the state-transition matrix of the homogeneous system

$$\dot{\mathbf{z}}(t) = \mathbf{A}(t) \mathbf{z}(t), \quad \mathbf{z}(t_0) = \mathbf{z}_0, \quad \forall t \geq t_0.$$

The variation in the active terminal constraint values \mathbf{T}^a caused by the input variation (6) is given by the Gâteaux derivative [4] –provided it exists– of \mathbf{T}^a with respect to $\boldsymbol{\xi}^u$ at \mathbf{u}^* :

$$\delta \mathbf{T}^a(\mathbf{u}^*; \boldsymbol{\xi}^u) := \left. \frac{\partial}{\partial \eta} \mathbf{T}^a(t_f, \tilde{\mathbf{x}}(t_f; \eta), \boldsymbol{\theta}_0) \right|_{\eta=0} = \mathbf{T}_{\mathbf{x}}^a[t_f] \boldsymbol{\xi}^x(t_f),$$

which using (8) becomes

$$\delta \mathbf{T}^a(\mathbf{u}^*; \boldsymbol{\xi}^u) = \mathbf{T}_{\mathbf{x}}^a[t_f] \int_{t_0}^{t_f} \boldsymbol{\Phi}^{\mathbf{f}_x}(t_f, s) \mathbf{f}_{\mathbf{u}}[s] \boldsymbol{\xi}^u(s) ds. \quad (9)$$

If the input variation $\boldsymbol{\xi}^u$ does not cause any change in the value of the terminal constraints, we have

$$\mathcal{D} \boldsymbol{\xi}^u = \mathbf{0}, \quad (10)$$

³The following compact notation is used subsequently:

$$y[t] := y(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), \boldsymbol{\rho}^*, \boldsymbol{\theta}_0).$$

¹The following notation is used throughout the paper: $\mathcal{I}_n := \{1, \dots, n\}$.

²The notation $\mathbf{f}_{\mathbf{z}}$ is used for the Jacobian matrix of the vector function \mathbf{f} with respect to the vector \mathbf{z} .

where $\mathcal{D} : \mathcal{C}[t_0, t_f]^{n_u} \rightarrow \mathbb{R}^{n_{T^a}}$ defined by

$$\mathcal{D}\xi := \mathbf{T}_x^a[t_f] \int_{t_0}^{t_f} \Phi^{\mathbf{f}^x}(t_f, s) \mathbf{f}_u[s] \xi(s) ds$$

is a linear operator. Equation (10) is a linear Fredholm integral equation of the first kind in ξ^u [5]. Hence, the input variation functions ξ^u that do not cause any change in the value of the active terminal constraints are those satisfying (10).

We are now ready to define the sensitivity-seeking (SS) directions.

Definition 1 (Sensitivity-seeking Directions): A function ξ^u in the input function space $\mathcal{C}[t_0, t_f]^{n_u}$ along which an infinitesimal input variation $\eta \xi^u$ does not modify the values of the active terminal constraints $\mathbf{T}^a(t_f, \mathbf{x}^*(t_f), \boldsymbol{\theta}_0)$, in the sense of (10), is called a sensitivity-seeking direction in $\mathcal{C}[t_0, t_f]^{n_u}$.

The foregoing discussion makes it clear that any solution of (10) yields a SS direction and so does any linear combination of solutions of (10). Let the set of all solutions to (10) be denoted by \mathcal{V}^s ,

$$\mathcal{V}^s := \{ \mathbf{v} \in \mathcal{C}[t_0, t_f]^{n_u} \mid \mathcal{D}\mathbf{v} = \mathbf{0} \}.$$

Note that \mathcal{V}^s is a linear subspace of $\mathcal{C}[t_0, t_f]^{n_u}$. Since \mathcal{V}^s is the span of the SS directions, it shall be referred to as the *sensitivity-seeking subspace* of the input function space subsequently.

Next, a constraint-seeking (CS) direction is defined as one that is *orthogonal* to the sensitivity-seeking subspace.

Definition 2 (Constraint-seeking Directions): A function $\xi^u \in \mathcal{C}[t_0, t_f]^{n_u}$ is called a constraint-seeking direction for the optimal control problem $OC(\boldsymbol{\theta}_0)$ at \mathbf{u}^* if ξ^u is orthogonal to \mathcal{V}^s , i.e.,

$$0 = \langle \xi^u, \xi \rangle, \quad \forall \xi \in \mathcal{V}^s,$$

where $\langle \cdot, \cdot \rangle$ stands for any inner product on $\mathcal{C}[t_0, t_f]^{n_u}$.

Let us denote by \mathcal{V}^c the set

$$\mathcal{V}^c := \{ \mathbf{v} \in \mathcal{C}[t_0, t_f]^{n_u} \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0, \quad \forall \mathbf{w} \in \mathcal{V}^s \}.$$

Note that, by the linearity property of the inner product, \mathcal{V}^c is also a linear subspace of $\mathcal{C}[t_0, t_f]^{n_u}$. Since \mathcal{V}^c is the span of CS directions, it shall be referred to as the *constraint-seeking subspace* of the input function space subsequently.

Lemma 3: No non-zero $\mathbf{v}_c \in \mathcal{V}^c$ satisfies (10), i.e.,

$$\mathcal{V}^s \cap \mathcal{V}^c = \{ \mathbf{0} \}.$$

Proof: Let $\xi \in \mathcal{V}^s \cap \mathcal{V}^c$. By construction, we have $\langle \xi, \xi \rangle = 0$, which by the elementary properties of an inner product implies $\xi = \mathbf{0}$. ■

At this point, it is easy to define selective input adaptation along each of the set of directions defined above.

Definition 4 (Selective Input Adaptation): Adaptation of the nominal optimal inputs according to (6) along the nonzero direction $\xi^u \in \mathcal{V}^s$ is called *selective input adaptation along a SS direction*. Similarly, an adaptation along a nonzero direction $\xi^u \in \mathcal{V}^c$ is called *selective input adaptation along a CS direction*.

Note that, because the CS and SS directions are tied to the nominal optimal solution, no variation in the parameters $\boldsymbol{\theta}$ is involved in the foregoing definitions.

IV. EFFECT OF SELECTIVE INPUT ADAPTATION ON COST

This section deals with the scenario of a change in parameters from $\boldsymbol{\theta}_0$ to $\tilde{\boldsymbol{\theta}}(\eta) := \boldsymbol{\theta}_0 + \eta \xi^\theta$, with $|\eta| \ll 1$.

Suppose that one wishes to avoid repeating the whole solution procedure to compute the modified optimal inputs $\tilde{\mathbf{u}}^*$. Either one of two options are possible:

- 1) *No Input Adaptation:* The nominal optimal inputs \mathbf{u}^* are applied ‘as is’ to the perturbed system. Let the pair of perturbed states and resulting cost be denoted by $(\tilde{\mathbf{x}}(t), \tilde{J})$. Thus, $(\tilde{\mathbf{x}}(t), \mathbf{u}^*(t))$ satisfies (1) for $\tilde{\boldsymbol{\theta}}$. Because of the continuous differentiability of \mathbf{f} with respect to \mathbf{x} and $\boldsymbol{\theta}$, $\tilde{\mathbf{x}}(t)$ has a first-order approximation around $\mathbf{x}^*(t)$ as

$$\tilde{\mathbf{x}}(t; \eta) = \mathbf{x}^*(t) + \eta \dot{\xi}^{\tilde{\mathbf{x}}}(t) + O(\eta^2).$$

- 2) *Selective Input Adaptation:* The nominal optimal inputs are adapted along a general direction $\xi^u(t) \in \mathcal{C}[t_0, t_f]$ and the resulting inputs (6) are then applied to the perturbed system. Let the pair of perturbed states and resulting cost be denoted by $(\tilde{\mathbf{x}}(t), \tilde{J})$, respectively. Thus, $(\tilde{\mathbf{x}}(t), \tilde{\mathbf{u}}(t))$ satisfies (1) for $\tilde{\boldsymbol{\theta}}$. Because of the continuous differentiability of \mathbf{f} with respect to \mathbf{x} , \mathbf{u} and $\boldsymbol{\theta}$, $\tilde{\mathbf{x}}(t)$ also has a first-order approximation around $\mathbf{x}^*(t)$ as

$$\tilde{\mathbf{x}}(t; \eta) = \mathbf{x}^*(t) + \eta \dot{\xi}^{\tilde{\mathbf{x}}}(t) + O(\eta^2).$$

Subscript *s* or *c* will be added to various notation when the direction of input adaptation $\xi^u(t)$ under consideration is a SS or a CS direction, respectively.

Evidently, both of the above options will result in sub-optimal process operation, although Option 2 can be expected to perform better under judicious choice of the input adaptation directions. The objective here is to compare the cost variations $\tilde{J}_s - \tilde{J}$ and $\tilde{J}_c - \tilde{J}$.

Following common practice in optimal control theory [6], the cost functional J is augmented as

$$J^a := \psi(t_f, \mathbf{x}(t_f), \boldsymbol{\theta}) + \int_{t_0}^{t_f} \phi(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}) dt + \int_{t_0}^{t_f} \boldsymbol{\pi}(t)^T [\mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}) - \dot{\mathbf{x}}(t)] dt,$$

for some multiplier functions $\boldsymbol{\pi} \in \mathcal{C}^1[t_0, t_f]^{n_x}$. It is clear that $J^a = J$ for any $\boldsymbol{\pi}(t) \in \mathcal{C}^1[t_0, t_f]^{n_x}$ provided that the pair $(\mathbf{x}(t), \mathbf{u}(t))$ satisfies (1) for $\boldsymbol{\theta}$, in which case minimizing J with respect to \mathbf{u} becomes *equivalent* to minimizing J^a with respect to \mathbf{u} . Using integration by parts, the expression of J^a can be rearranged as follows:

$$J^a = \psi(t_f, \mathbf{x}(t_f), \boldsymbol{\theta}) - \boldsymbol{\pi}(t_f)^T \mathbf{x}(t_f) + \boldsymbol{\pi}(t_0)^T \mathbf{h}(\boldsymbol{\theta}) + \int_{t_0}^{t_f} [\phi(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}) + \boldsymbol{\pi}(t)^T \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta})] dt + \int_{t_0}^{t_f} \dot{\boldsymbol{\pi}}(t)^T \mathbf{x}(t) dt. \quad (11)$$

Since both pairs $(\tilde{\mathbf{x}}(t), \tilde{\mathbf{u}}(t))$ and $(\hat{\mathbf{x}}(t), \mathbf{u}^*(t))$ satisfy (1) for $\tilde{\theta}$,⁴

$$\begin{aligned} \tilde{J}^a - \hat{J}^a &= \tilde{\psi}[t_f] - \hat{\psi}[t_f] - \boldsymbol{\pi}(t_f)^T \tilde{\mathbf{x}}(t_f) + \boldsymbol{\pi}(t_f)^T \hat{\mathbf{x}}(t_f) \\ &+ \int_{t_0}^{t_f} \left(\tilde{\phi}[t] - \hat{\phi}[t] \right) dt \\ &+ \int_{t_0}^{t_f} \left(\boldsymbol{\pi}(t)^T \tilde{\mathbf{f}}[t] - \boldsymbol{\pi}(t)^T \hat{\mathbf{f}}[t] \right) dt \\ &+ \int_{t_0}^{t_f} \left(\tilde{\boldsymbol{\pi}}(t)^T \tilde{\mathbf{x}}(t) - \tilde{\boldsymbol{\pi}}(t)^T \hat{\mathbf{x}}(t) \right) dt. \end{aligned} \quad (12)$$

Taylor expansion around $(\hat{\mathbf{x}}(t), \mathbf{u}^*(t))$ and rearrangement of the various terms in (12) leads to:

$$\begin{aligned} \tilde{J}^a - \hat{J}^a &= \eta \left\{ \left(\hat{\psi}_{\mathbf{x}}[t_f]^T - \boldsymbol{\pi}(t_f)^T \right) \boldsymbol{\zeta}^{\mathbf{x}}(t_f) \right. \\ &+ \int_{t_0}^{t_f} \left(\hat{\phi}_{\mathbf{x}}[t]^T + \boldsymbol{\pi}(t)^T \hat{\mathbf{f}}_{\mathbf{x}}[t] + \tilde{\boldsymbol{\pi}}(t)^T \right) \boldsymbol{\zeta}^{\mathbf{x}}(t) dt \\ &+ \left. \int_{t_0}^{t_f} \left(\hat{\phi}_{\mathbf{u}}[t]^T + \boldsymbol{\pi}(t)^T \hat{\mathbf{f}}_{\mathbf{u}}[t] \right) \boldsymbol{\xi}^{\mathbf{u}}(t) dt \right\} + O(\eta^2), \end{aligned}$$

where

$$\boldsymbol{\zeta}^{\mathbf{x}}(t) := \boldsymbol{\xi}^{\tilde{\mathbf{x}}}(t) - \boldsymbol{\xi}^{\hat{\mathbf{x}}}(t), \quad \forall t \in [t_0, t_f].$$

With the following choice of $\boldsymbol{\pi}$,

$$\begin{aligned} \dot{\hat{\boldsymbol{\pi}}}(t) &= -\hat{\mathbf{f}}_{\mathbf{x}}[t]^T \hat{\boldsymbol{\pi}}(t) - \hat{\phi}_{\mathbf{x}}[t], \\ \hat{\boldsymbol{\pi}}(t_f) &= \hat{\psi}_{\mathbf{x}}[t_f], \end{aligned} \quad (13)$$

the cost difference further reduces to

$$\tilde{J}^a - \hat{J}^a = \eta \int_{t_0}^{t_f} \left(\hat{\phi}_{\mathbf{u}}[t]^T + \hat{\boldsymbol{\pi}}(t)^T \hat{\mathbf{f}}_{\mathbf{u}}[t] \right) \boldsymbol{\xi}^{\mathbf{u}}(t) dt + O(\eta^2).$$

Note that the inhomogeneous linear differential equations (13) can always be solved to obtain a unique $\hat{\boldsymbol{\pi}}(t)$. In turn, $\hat{\phi}_{\mathbf{u}}[t]$ and $\hat{\mathbf{f}}_{\mathbf{u}}[t]$ can be expanded around $(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\theta}_0)$ to give

$$\tilde{J}^a - \hat{J}^a = \eta \int_{t_0}^{t_f} \left(\phi_{\mathbf{u}}[t]^T + \hat{\boldsymbol{\pi}}(t)^T \mathbf{f}_{\mathbf{u}}[t] \right) \boldsymbol{\xi}^{\mathbf{u}}(t) dt + O(\eta^2).$$

Hence, a first-order approximation of \tilde{J}^a is $\hat{J}^a + \eta \delta J$, with

$$\delta J := \int_{t_0}^{t_f} \left(\phi_{\mathbf{u}}[t]^T + \hat{\boldsymbol{\pi}}(t)^T \mathbf{f}_{\mathbf{u}}[t] \right) \boldsymbol{\xi}^{\mathbf{u}}(t) dt. \quad (14)$$

Next, utilizing the fact that (4) holds along the nominal optimal trajectory,

$$\mathbf{0} = \mathcal{H}_{\mathbf{u}}[t] = \phi_{\mathbf{u}}[t] + \mathbf{f}_{\mathbf{u}}[t]^T \boldsymbol{\lambda}^*(t), \quad \text{a.e. in } [t_0, t_f],$$

δJ can be rewritten as

$$\delta J = \int_{t_0}^{t_f} \beta(t)^T \mathbf{f}_{\mathbf{u}}[t] \boldsymbol{\xi}^{\mathbf{u}}(t) dt, \quad (15)$$

⁴We introduce the additional notations:

$$\tilde{y}[t] := y(t, \tilde{\mathbf{x}}(t), \mathbf{u}^*(t), \tilde{\boldsymbol{\theta}}), \quad \text{and} \quad \hat{y}[t] := y(t, \hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t), \tilde{\boldsymbol{\theta}}).$$

where $\beta(t)$ is the first-order approximation of $\hat{\boldsymbol{\pi}}(t) - \boldsymbol{\lambda}^*(t)$, i.e.,

$$\hat{\boldsymbol{\pi}}(t) - \boldsymbol{\lambda}^*(t) = \beta(t) + \eta \boldsymbol{\xi}^{\boldsymbol{\pi}}(t) + O(\eta^2),$$

and is given by

$$\beta(t) = -\boldsymbol{\Phi}^{\mathbf{f}_{\mathbf{x}}}(t_f, t)^T \mathbf{T}_{\mathbf{x}}^a[t_f]^T \boldsymbol{\rho}_a, \quad \forall t \in [t_0, t_f]. \quad (16)$$

Finally, combining (15) and (16) gives:

$$\delta J = -\boldsymbol{\rho}_a^T \mathbf{T}_{\mathbf{x}}^a[t_f] \int_{t_0}^{t_f} \boldsymbol{\Phi}^{\mathbf{f}_{\mathbf{x}}}(t_f, t) \mathbf{f}_{\mathbf{u}}[t] \boldsymbol{\xi}^{\mathbf{u}}(t) dt. \quad (17)$$

Now we are ready to state the following theorem:

Theorem 5: Consider parametric variations of the form $\tilde{\boldsymbol{\theta}}(\eta) := \boldsymbol{\theta}_0 + \eta \boldsymbol{\xi}^{\boldsymbol{\theta}}$, with $|\eta| \ll 1$. The cost variation by making selective input adaptation along any nonzero SS direction $\boldsymbol{\xi}_s^{\mathbf{u}} \in \mathcal{V}^s$ is $O(\eta^2)$, whereas that by making selective input adaptation along any nonzero CS direction $\boldsymbol{\xi}_c^{\mathbf{u}} \in \mathcal{V}^c$ is $O(\eta)$.

Proof: By Definition 1, $\boldsymbol{\xi}_s^{\mathbf{u}}$ satisfies (10), and from (17),

$$\delta J_s = -\boldsymbol{\rho}_a^T \mathbf{T}_{\mathbf{x}}^a[t_f] \int_{t_0}^{t_f} \boldsymbol{\Phi}^{\mathbf{f}_{\mathbf{x}}}(t_f, t) \mathbf{f}_{\mathbf{u}}[t] \boldsymbol{\xi}_s^{\mathbf{u}}(t) dt = 0.$$

Therefore, $\tilde{J}_s^a - \hat{J}^a$ is $O(\eta^2)$. On the other hand, no non-zero direction in \mathcal{V}^c satisfies (10) by Lemma 3. Hence, from (17), $\delta J_c = -\boldsymbol{\rho}_a^T \{\text{non-zero vector}\} \neq 0$, as strict complementarity condition holds at the nominal optimal solution of $(\text{OC}(\boldsymbol{\theta}))$. Therefore, $\tilde{J}_c^a - \hat{J}^a$ is $O(\eta)$. ■

The main implication of Theorem 5 is that, for small parametric variations, adapting the inputs along the constraint-seeking directions has the greatest impact on the performance of the perturbed system. Note, however, that no restriction has been considered so far regarding the choice of the input adaptation directions $\boldsymbol{\xi}_c^{\mathbf{u}}$. A judicious choice of $\boldsymbol{\xi}_c^{\mathbf{u}}$ will permit a substantial *cost improvement*, while a poor choice of $\boldsymbol{\xi}_c^{\mathbf{u}}$ could potentially lead to worsening the performance of the adapted system, even with respect to the no-input-adaptation scenario (Option 1). Special care must therefore be taken when selecting the input-adaptation directions. A numerical procedure for computing sensitivity- and constraint-seeking directions and a way of choosing the input adaptation directions are discussed in the following section.

V. A NUMERICAL PROCEDURE TO COMPUTE SENSITIVITY- AND CONSTRAINT-SEEKING DIRECTIONS

This section proposes a numerical procedure to compute sensitivity- and constraint-seeking directions.

Let $\boldsymbol{\xi}^{\mathbf{u}}$ denote a given direction in the input function space. We would like to compute the SS and CS directions $\boldsymbol{\xi}_s^{\mathbf{u}}$ and $\boldsymbol{\xi}_c^{\mathbf{u}}$ obtained by projecting $\boldsymbol{\xi}^{\mathbf{u}}$ on \mathcal{V}_s and \mathcal{V}_c , respectively.

To avoid the difficulty of computing projections on the infinite-dimensional function spaces \mathcal{V}_c and \mathcal{V}_s , we propose to proceed by approximating the optimal control problem by a nonlinear programming problem (NLP) as follows:

- 1) Approximate the input profiles $\mathbf{u}(t)$ using a control vector parameterization (e.g., piecewise constant or affine) in terms of n parameters, the vector of which

will be denoted by ω . Thus, we have the following expression which relates ω to $\mathbf{u}(t)$:

$$\mathbf{u}(t) = \mathcal{U}(t, \omega), \quad \forall t \in [t_0, t_f]. \quad (18)$$

- 2) Transform the optimal control problem into an NLP in terms of the decision variables ω . Note that the terminal constraint in the optimal control problem will be transformed into a - typically nonlinear - constraint in ω .
- 3) Solve the resulting NLP numerically to obtain the optimal values ω^* and find the set of active constraints \mathbf{G} of the NLP at ω^* .
- 4) From the singular value decomposition of \mathbf{G}_ω at ω^* , find the orthogonal matrices \mathbf{V}_c and \mathbf{V}_s that define the *CS and SS directions, respectively, of the NLP problem* [7].
- 5) Compute the orthogonal projections of the vector ξ^ω on the column space of \mathbf{V}_c and \mathbf{V}_s , respectively:

$$\begin{aligned} \xi_c^\omega &= \mathbf{V}_c \mathbf{V}_c^T \xi^\omega, \\ \xi_s^\omega &= \mathbf{V}_s \mathbf{V}_s^T \xi^\omega. \end{aligned} \quad (19)$$

- 6) ξ_c^ω and ξ_s^ω are the approximations of the desired profiles ξ_c^u and ξ_s^u , respectively, under the same parameterization as used in Step 1, i.e.,

$$\begin{aligned} \xi_c^u(t) &= \mathcal{U}(t, \xi_c^\omega), \quad \forall t \in [t_0, t_f], \\ \xi_s^u(t) &= \mathcal{U}(t, \xi_s^\omega), \quad \forall t \in [t_0, t_f]. \end{aligned}$$

Naturally, we can expect the approach to yield better approximations of the desired directions ξ_c^u and ξ_s^u as the number of parameters n increases.

In case of small parametric variations around θ_0 , it is possible to specialize the choice of the input variation $\xi^u(t)$ to the first-order input variation $\xi^{u^*}(t)$ [1]. Note that, in the present approach, it is not necessary to compute $\xi^{u^*}(t)$ using the sensitivity analysis of the optimal control problem. Indeed, if ξ^{ω^*} denotes the parameterization of $\xi^{u^*}(t)$ of the type chosen in Step 1 above, then ξ^{ω^*} can be computed from the sensitivity analysis of the (resulting) NLP problem [7], [8], [9]. Steps 3 to 6 will then yield the specific input adaptation directions $\xi_c^{u^*}(t) \in \mathcal{V}_c$ and $\xi_s^{u^*}(t) \in \mathcal{V}_s$, respectively.

VI. ILLUSTRATIVE EXAMPLE

Consider the following parametric optimal control problem corresponding to the optimization of a chemical reactor. There is one input variable, the reactor temperature, and one terminal constraint [10]:

$$\begin{aligned} & \max_{u(t), 0 \leq t \leq t_f} x_B(t_f) \\ \text{s. t.} \quad & \dot{x}_A(t) = -k_1(u(t))x_A(t) \\ & \dot{x}_B(t) = k_1(u(t))x_A(t) - k_2(u(t))x_B(t) \\ & k_1(u(t)) = k_1^\circ \exp\left(-\frac{E_1}{u(t)}\right) \\ & k_2(u(t)) = k_2^\circ \exp\left(-\frac{E_2}{u(t)}\right) \end{aligned} \quad (20)$$

$$\begin{aligned} x_A(0) &= 0.53, \quad x_B(0) = 0.43 \\ x_A(t_f) - 0.1(2 - \theta) &\leq 0, \end{aligned}$$

where θ denotes the uncertain system parameter, with nominal value $\theta_0 = 1$.

The values of all constants in the equations above are given in Table I.

TABLE I
CONSTANTS

Constants	Values
k_1°	0.535×10^{11}
k_2°	0.461×10^{18}
E_1	9×10^3
E_2	15×10^3
t_f	8

Following the procedure outlined in Section V, a piecewise-constant input parameterization over $n = 90$ equal-length stages over $[0, t_f]$ is considered. Figure 1 shows the nominal optimal solution of (20) reconstructed using the solution of the resulting NLP.

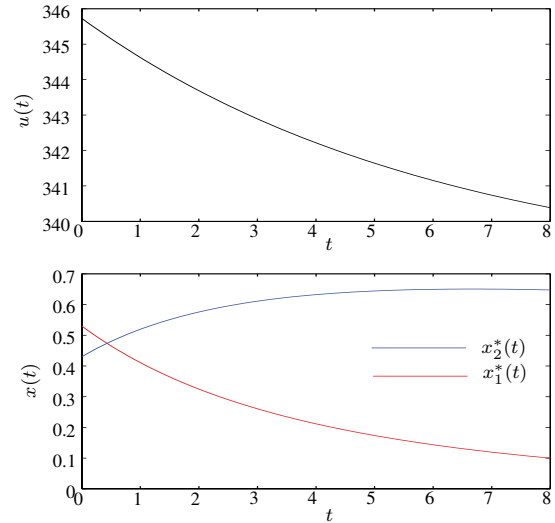


Fig. 1. Nominal optimal input profile (top plot) and corresponding state trajectories (bottom plot).

Next, ξ^{ω^*} is computed from the sensitivity analysis of the NLP problem for $\xi^\theta = 1$. Using ξ^{ω^*} in (19) yields $\xi_c^{\omega^*}$ and $\xi_s^{\omega^*}$. These three vectors are used to reconstruct approximations for the first-order input variation $\xi^{u^*}(t)$ and corresponding CS and SS directions $\xi_c^{u^*}(t)$ and $\xi_s^{u^*}(t)$, respectively. The latter three functions are shown in Figure 2.

Finally, the performances of the following strategies are compared for 5% and 10% variation in the parameter θ , i.e., $\eta = 0.05$ and 0.1:

- *No input adaptation* : $\tilde{u}(t) = u^*(t)$.
- *Full input adaptation* : $\tilde{u}(t) = u^*(t) + \eta \xi^{u^*}(t)$.
- *Selective input adaptation in the CS direction only* : $\tilde{u}(t) = u^*(t) + \eta \xi_c^{u^*}(t)$.

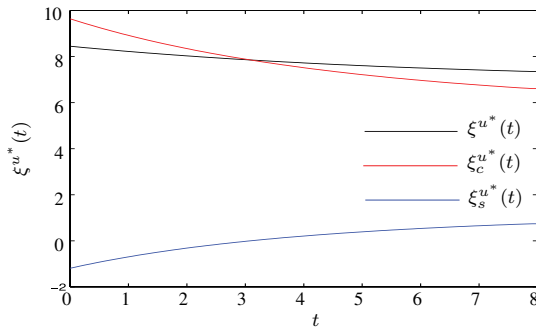


Fig. 2. Estimates of the first-order input variation $\xi^{u^*}(t)$ and the corresponding CS and SS directions $\xi_c^{u^*}(t)$ and $\xi_s^{u^*}(t)$.

The results are given in Table II.

TABLE II
RESULTS OF INPUT ADAPTATION STRATEGIES

Strategy	J	$T(t_f, \mathbf{x}(t_f))$
$\eta = 0.05$		
No adaptation	0.648	500×10^{-5}
CS-selective adaptation	0.642	6×10^{-5}
Full adaptation	0.642	6×10^{-5}
Perturbed optimal	0.642	0
$\eta = 0.1$		
No adaptation	0.648	1000×10^{-5}
CS-selective adaptation	0.636	23×10^{-5}
Full adaptation	0.636	23×10^{-5}
Perturbed optimal	0.635	0

In both perturbation scenarios, the performance of full and CS-selective input adaptation are almost identical, thereby validating the principle of selective input adaptation. Note also that the terminal constraint violation by the two adaptation strategies is negligible compared to that due to the strategy of no adaptation in both perturbation scenarios.

There is degradation, in terms of the value of terminal constraint, in the performance of both input adaptation strategies, when the magnitude of the parametric variation (η) increases from 5% to 10%. This latter effect results from the first-order approximations becoming less accurate as the magnitude of the parametric variation increases. The degradation effect on terminal constraint violation is very pronounced in case of no input adaptation.

The results also make it evident that it may become necessary to back-off the terminal constraints in practical implementation of input adaptation strategies.

VII. CONCLUSIONS

The complexity of optimal control problems makes those methods that do not require recomputing the exact solution to be much more tractable and thus highly desirable.

For problems involving terminal constraints, it is demonstrated that the input function space can be split into two distinct sets of directions, based on whether an infinitesimal

input variation along these directions modifies the terminal constraints or not; the criterion for separation turns out to be a linear Fredholm integral equation of the first kind. These two orthogonal sets of directions are termed constraint-seeking and sensitivity-seeking directions, respectively.

In order to mitigate the effect of parametric variations, while keeping the adaptation problem tractable, selective input adaptation along either the SS directions or the CS directions is considered. It is shown that the cost variation by making the latter adaptation is $O(\eta)$, whereas it is $O(\eta^2)$ with the former one. This paves the way for the development of selective input adaptation schemes which will greatly reduce, if not eliminate, the need for reoptimizing a system in the presence of parametric variations.

A possible application of these results is in the field of a recently developed methodology for constrained optimal control problems called NCO tracking [11], in which parts of the input profiles can be adapted selectively. Hence, prioritization of selective-adaptation strategies is of paramount importance for developing practical NCO-tracking controllers.

Extensions of these results to problems involving discontinuous $\mathbf{u}^*(t)$, problems having non-regular Hamiltonians and singular arcs as well as problems having a combination of terminal-, mixed control-state- and pure state-constraints will be addressed in future work.

ACKNOWLEDGMENTS

Financial support from the Swiss National Science Foundation (SNSF) under Grant #200020-113292 is gratefully acknowledged.

REFERENCES

- [1] S. Deshpande, B. Chachuat, and D. Bonvin, "Parametric Sensitivity of Path-Constrained Optimal Control: Towards Selective Input Adaptation," in *American Control Conference 2009*, 2009, pp. 349–354.
- [2] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM Review*, vol. 37, no. 2, pp. 181–218, 1995.
- [3] W. J. Rugh, *Linear System Theory*. Englewood Cliffs, New Jersey: Prentice Hall, 1993.
- [4] A. Wouk, *A Course of Applied Functional Analysis*. New York: John Wiley & Sons, Inc., 1979.
- [5] G. M. Wing, *A Primer on Integral Equations of the First Kind*. Philadelphia: SIAM, 1991.
- [6] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. Washington DC: Hemisphere, 1975.
- [7] B. Chachuat, A. Marchetti, and D. Bonvin, "Process Optimization via Constraints Adaptation," *Journal of Process Control*, vol. 18, no. 3-4, pp. 244–257, 2008.
- [8] A. V. Fiacco, "Sensitivity analysis for nonlinear programming using penalty methods," *Mathematical Programming*, vol. 10, no. 1, pp. 287–311, 12 1976.
- [9] —, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, ser. Mathematics in Science and Engineering. Academic Press, 1983.
- [10] W. F. Ramirez, *Process Control and Identification*. Academic Press, 1994.
- [11] B. Srinivasan and D. Bonvin, "Real-Time Optimization of Batch Processes via Tracking of Necessary Conditions of Optimality," *Ind Eng Chem Res*, vol. 46, no. 2, pp. 492–504, 2007.