Static Optimization based on Output Feedback

S. Gros, B. Srinivasan*, and D. Bonvin

Laboratoire d'Automatique

École Polytechnique Fédérale de Lausanne

CH-1015 Lausanne, Switzerland

January 21, 2007

Abstract

In the framework of process optimization, the use of measurements to compensate the effect of uncertainty has re-emerged as an active area of research. One of the ideas therein is to adapt the inputs in order to track the active constraints and push certain sensitivities to zero. In perturbation-based optimization, the sensitivities are evaluated by perturbation of the inputs and measurement of the cost function, which can be experimentally time consuming. However, since more measurements (typically the outputs) than just the cost function are available, the idea developed in this paper is to incorporate the outputs in a measurement-based optimization framework. This is done using an extension to the neighboring-extremal scheme for the case of output measurements. If measurement noise can be neglected, the approach is shown to converge to the optimum in at most two input updates. The effect of measurement noise is

*Current address: Department of Chemical Engineering, Ecole Polytechnique Montreal, Montreal, Canada

also investigated. The strength of neighboring-extremal output feedback for optimization is illustrated on a continuous chemical reactor example.

Keywords: Static optimization, Measurement-based optimization, Gradient estimation, NCO tracking, Neighboring extremals, Output feedback, Measurement noise.

1 Introduction

The optimization of dynamic systems amidst uncertainty has re-gained popularity recently (Kristic and Wang, 2000; Skogestad, 2000; Zang *et al.*, 2001; Srinivasan *et al.*, 2003) after some initial work in the 1950s (Morosanov, 1957; Ostrovskii, 1957; Pervozvanskii, 1960). This renewed interest is mainly due to advances in instrumentation and thus to the availability of measurements. In this context, two optimization classes need to be distinguished: (i) Optimization of a steady-state operating point (Marlin and Hrymak, 1996; Kristic and Wang, 2000), and (ii) optimization of input profiles (Srinivasan *et al.*, 2003; Kadam and Marquardt, 2004). The former can be treated as a static optimization problem (dynamic systems operated at steady state), while the latter requires dynamic optimization tools. It can be argued that the former has considerable industrial impact due to the large equipment size and production volume associated with these processes. Examples of continuous processes are numerous in the process industry and include continuous chemical production and mineral processing.

When production is performed at steady state, it is critical that the system be operated as closely as possible to the optimal operating point. Standard optimization tools rely on a process model, which, for industrial applications, are often inaccurate or incomplete. Model mismatch is usually the result of simplifications and process variations (Eaton and Rawlings, 1990; Abel and Marquardt, 1998; Ruppen *et al.*, 1995). Hence, the optimal setpoints computed from the available models are typically not optimal for the reality.

On the other hand, measurement-based optimization uses appropriate measurements to compensate the effect of uncertainty. The measurements are used to either: (i) Adapt the parameters of a process model and re-optimize it (*explicit optimization*) (Marlin and Hrymak, 1996; Zang *et al.*, 2001; Kadam *et al.*, 2003), or (ii) directly adapt the inputs (*implicit optimization*) (Kristic and Wang, 2000; Skogestad, 2000; Srinivasan *et al.*, 2003). Furthermore, in static implicit optimization, it is possible to distinguish between three types of techniques:

1. Zeroth-order methods – In techniques labeled *evolutionary optimization* (Box and Draper, 1987), a simplex-type algorithm is used to approach the optimum. The cost function is measured experimentally for every combination of operating conditions.

2. First-order methods – In techniques labeled *extremum-seeking control* (Kristic and Wang, 2000), the gradients are estimated experimentally using sinusoidal excitation. The excitation frequency needs to be sufficiently small for a time-scale separation between the system dynamics and the excitation frequency to exist. Like the techniques of the first type, this scheme also uses only the measurement of the cost function.

3. Reformulation methods – In techniques such as *self-optimizing control* (Skogestad, 2000) or *NCO tracking* (François *et al.*, 2005), the optimization problem is recast as a problem of choosing outputs whose optimal values are approximately invariant to uncertainty. The output values vary with uncertainty, but they are brought back to their invariant setpoints using measurements. These schemes use output information rather than simply the measurement of the cost function.

This paper proposes a first-order method that uses output information to compute the gradients in one go, i.e. without having to rely on several measurements at different operating points as this is the case when only the cost function is measured. Thus, the number of iterations required to reach the optimum is considerably smaller than with a perturbation-based approach. Gradient computation is done implicitly using the neighboring-extremal (NE) approach that is well known in optimal control (Bryson and Ho, 1975). The NE approach typically relies on state feedback. In this paper, in order to be attractive from an implementation point of view, the NE approach is modified to use only the available (output) measurements.

The paper is organized as follows. Section 2 formulates the optimization problem and introduces the neighboring-extremal and gradient-based methods to optimization. Section 3 details the output NE scheme and proves its local convergence. In addition, the effect of measurement noise on the performance of the update scheme is studied. The optimization of an isothermal continuous reactor is considered in Section 4, while Section 5 provides conclusions.

2 Preliminaries

2.1 Optimization Problem

Steady-state optimization consists of minimizing a given cost function under equality and inequality constraints. At the optimum, some of the inequality constraints are active. A standard assumption is that the set of active constraints does not change with uncertainty that includes model mismatch and process disturbances. In such a case, these active constraints can be kept active using simple controllers, which in turn removes certain degrees of freedom from the optimization problem. Thus, a problem without inequality constraints and a smaller set of decision variables can be formulated as is done in most approaches (Skogestad, 2000; Kristic and Wang, 2000; Zang *et al.*, 2001).

The general formulation of a static optimization problem without inequality constraints is considered:

$$\min_u \phi(x, u, \theta) \tag{1}$$

$$s.t. \qquad F(x, u, \theta) = 0 \tag{2}$$

where ϕ is the smooth scalar cost function to be minimized, u the m-dimensional vector of inputs, x the n-dimensional vector of states, θ the n_θ-dimensional vector of parameters, and F the n-dimensional vector of algebraic equations that describe the dynamic system at steady state. Note that the system equations F can be solved for x and substituted into ϕ. Such a simplification is purposely not done here since the information on x will be used explicitly to compute the sensitivities.

2.2 Optimality Conditions

Introducing the Lagragian $L(x, u, \theta, \lambda) = \phi(x, u, \theta) + \lambda^T F(x, u, \theta)$, where λ are the adjoint variables, the necessary conditions of optimality (NCO) for Problem (1)-(2) read:

$$L_u = \phi_u + \lambda^T F_u = 0 \tag{3}$$

$$L_x = \phi_x + \lambda^T F_x = 0 \tag{4}$$

$$L_\lambda = F^T = 0 \tag{5}$$

The notation $a_b = \frac{\partial a}{\partial b}$ is used in this paper. Assuming that F_x is invertible, the condition $L_x = 0$ defines the adjoint variables:

$$\lambda^T = -\phi_x F_x^{-1} \tag{6}$$

Using (6) in $L_u = 0$ gives:

$$L_u = \phi_u - \phi_x F_x^{-1} F_u = \frac{d\phi}{du} = 0 \tag{7}$$

with $\frac{d\phi}{du}$ being the total derivatives of the cost function that take into account both the direct effect of u and the effect of u through x. Hence, the total derivatives of the cost function with respect to u vanish at the optimum.

2.3 Input Update using Neighboring-extremal Control

Neighboring-extremal control attempts to maintain process optimality despite the presence of disturbances. It is based on the variations of the conditions of optimality.

Consider process operation around the nominal optimal operating point $F(x^*, u^*, \theta_{nom}) = 0$. Assume that the process is subject to the constant[1] parametric disturbance $\delta\theta = \theta - \theta_{nom}$, and let $\delta u_k = u_k - u^*$ be the input update computed at the discrete time instant k in an attempt to keep optimality. The deviations $\delta\theta$ and δu_k cause the state deviations $\delta x_k = x_k - x^*$. These deviations are linked together by the system equation (2), which, in linearized form, gives:

$$F_x \delta x_k + F_u \delta u_k + F_\theta \delta\theta = 0 \quad \forall k \tag{8}$$

For the process operation to remain optimal despite the parametric disturbance $\delta\theta$, the corrective action δu_{k+1} needs to bring the process to a new steady-state operating point at the time instant $k+1$. From (3)-(5), the first-order variations of the NCO read:

$$L_{ux} \delta x_{k+1} + L_{uu} \delta u_{k+1} + F_u^T \delta\lambda_{k+1} + L_{u\theta} \delta\theta = 0 \tag{9}$$

$$L_{xx} \delta x_{k+1} + L_{xu} \delta u_{k+1} + F_x^T \delta\lambda_{k+1} + L_{x\theta} \delta\theta = 0 \tag{10}$$

$$F_x \delta x_{k+1} + F_u \delta u_{k+1} + F_\theta \delta\theta = 0 \tag{11}$$

The three equations (9)-(11) can be solved for the optimal variations δx_{k+1}, δu_{k+1} and $\delta\lambda_{k+1}$ in

[1] A slowly time-varying disturbance can also be considered as long as the time-scale of the perturbation is large compared to the dynamics of the update scheme.

terms of the disturbance $\delta\theta$:

$$\begin{bmatrix} \delta x_{k+1} \\ \delta u_{k+1} \\ \delta \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} L_{ux} & L_{uu} & F_u^T \\ L_{xx} & L_{xu} & F_x^T \\ F_x & F_u & 0 \end{bmatrix}^{-1} \begin{bmatrix} L_{u\theta} \\ L_{x\theta} \\ F_\theta \end{bmatrix} \delta\theta \tag{12}$$

From (12), the update δu_{k+1} can be expressed analytically in terms of $\delta\theta$:

$$\delta u_{k+1} = K\delta\theta \tag{13}$$

with

$$\begin{aligned} K = & - (L_{uu} - L_{ux}F_x^{-1}F_u - F_u^T F_x^{-T} L_{xu} \\ & + F_u^T F_x^{-T} L_{xx} F_x^{-1} F_u)^{-1} (L_{u\theta} - L_{ux} F_x^{-1} F_\theta - F_u^T F_x^{-T} L_{x\theta} \\ & + F_u^T F_x^{-T} L_{xx} F_x^{-1} F_\theta) \end{aligned} \tag{14}$$

Since $\delta\theta$ is assumed constant, it can be estimated from the expression (8) at step k:

$$\delta\theta = -F_\theta^\dagger (F_x \delta x_k + F_u \delta u_k) \tag{15}$$

and used in the update law (13) at step $k+1$:

$$\delta u_{k+1} = \mathcal{K}^x \delta x_k + \mathcal{K}^u \delta u_k \tag{16}$$

with

$$\mathcal{K}^x = -K F_\theta^\dagger F_x \tag{17}$$

$$\mathcal{K}^u = -K F_\theta^\dagger F_u \tag{18}$$

Unfortunately, the update law (16) is based on full-state measurement δx_k, which is often not available in practice. In Section 3, it will be shown that the input update can be computed based on output measurements.

2.4 Gradient-based Approach to Optimization

Gradient-based approach schemes are based on the following input update equation:

$$u_{k+1} = u_k - \left(\frac{d^2\phi}{du^2}\right)_{u_k}^{-1} \frac{d\phi}{du}\bigg|_{u_k} \tag{19}$$

where u_k are the current inputs and u_{k+1} the inputs to be applied next. The gradient $\frac{d\phi}{du}\big|_{u_k}$ is evaluated experimentally via input perturbations around u_k. Since the Hessian $\left(\frac{d^2\phi}{du^2}\right)_{u_k}$ would be very expensive to compute experimentally, it is usually estimated from the model at the nominal operating point, thereby giving the following update law:

$$u_{k+1} = u_k - \left(\frac{d^2\phi}{du^2}\right)_{nom}^{-1} \frac{d\phi}{du}\bigg|_{u_k} \tag{20}$$

2.5 Link between Neigboring-extremal Control and Gradient-based Approach

It will be shown that the NE-control scheme (13), or equivalently (16), is a first-order approximation to the gradient-based law (20). First, a supporting lemma is presented that allows viewing the iterative scheme (20) in terms of deviations around the nominal operating point.

Lemma 1. *The gradient-based update scheme (20) can be written in terms of the deviations* δu_k *and* $\delta\theta$ *around the nominal optimal operating point.*

Proof. Subtracting u^* on both sides of (20) and introducing $\delta u_k = u_k - u^*$, the gradient-based strategy can be written as:

$$\delta u_{k+1} = \delta u_k - \left(\frac{d^2\phi}{du^2}\right)_{nom}^{-1} \frac{d\phi}{du}\bigg|_{u_k} \tag{21}$$

Taylor-series expansion of $\frac{d\phi}{du}\big|_{u_k}$ around the nominal optimal solution gives:

$$\frac{d\phi}{du}\bigg|_{u_k} = \left(\frac{d\phi}{du}\right)_{nom} + \left(\frac{d^2\phi}{du^2}\right)_{nom} \delta u_k + \left(\frac{d^2\phi}{du d\theta}\right)_{nom} \delta\theta + O(\delta u_k^2, \delta\theta^2) \tag{22}$$

Using (22) in (21) and considering the fact that $\left(\frac{d\phi}{du}\right)_{nom} = 0$ at the optimum, the gradient-based strategy becomes:

$$\delta u_{k+1} = -\left(\frac{d^2\phi}{du^2}\right)_{nom}^{-1} \left(\frac{d^2\phi}{dud\theta}\right)_{nom} \delta\theta + O(\delta u_k^2, \delta\theta^2) \tag{23}$$

□

Next, the proposition is established.

Proposition 1. *The NE-control scheme (13), or equivalently (16), is a first-order approximation to the gradient-based law (20).*

Proof. Lemma 1 indicates that the gradient-based law (20) can be written as (23), for which the terms $\left(\frac{d^2\phi}{du^2}\right)_{nom}$ and $\left(\frac{d^2\phi}{dud\theta}\right)_{nom}$ can be evaluated using the optimality conditions (7):

$$\left(\frac{d^2\phi}{du^2}\right)_{nom} = \frac{d}{du}L_u = L_{uu} + L_{ux}\frac{dx}{du} + F_u^T\frac{d\lambda}{du} \tag{24}$$

$$\left(\frac{d^2\phi}{dud\theta}\right)_{nom} = \frac{d}{d\theta}L_u = L_{u\theta} + L_{ux}\frac{dx}{d\theta} + F_u^T\frac{d\lambda}{d\theta} \tag{25}$$

The derivatives of x with respect to u and θ are obtained from the implicit function $F(x, u, \theta) = 0$ and read:

$$\frac{dx}{du} = -F_x^{-1}F_u \qquad \frac{dx}{d\theta} = -F_x^{-1}F_\theta \tag{26}$$

The adjoint variables λ are defined by the implicit functions $L_x(x, u, \theta, \lambda) = 0$, which are valid for all x, u and λ. Differentiation of $L_x = 0$ with respect to both u and λ gives:

$$\frac{d\lambda}{du} = -F_x^{-T}L_{xu} + F_x^{-T}L_{xx}F_x^{-1}F_u \tag{27}$$

$$\frac{d\lambda}{d\theta} = -F_x^{-T}L_{x\theta} + F_x^{-T}L_{xx}F_x^{-1}F_\theta \tag{28}$$

Using (26)-(28), equations (24) and (25) become:

$$\begin{aligned}\left(\frac{d^2\phi}{du^2}\right)_{nom} &= L_{uu} - L_{ux}F_x^{-1}F_u - F_u^T F_x^{-T} L_{xu} \\ &+ F_u^T F_x^{-T} L_{xx} F_x^{-1} F_u \quad (29)\\ \left(\frac{d^2\phi}{dud\theta}\right)_{nom} &= L_{u\theta} - L_{ux}F_x^{-1}F_\theta - F_u^T F_x^{-T} L_{x\theta} \\ &+ F_u^T F_x^{-T} L_{xx} F_x^{-1} F_\theta \quad (30)\end{aligned}$$

Using (29) and (30) in (14) gives:

$$K = -\left(\frac{d^2\phi}{du^2}\right)_{nom}^{-1} \left(\frac{d^2\phi}{dud\theta}\right)_{nom} \quad (31)$$

and the gradient-based update law (23) can be written as:

$$\delta u_{k+1} = K\delta\theta + O(\delta u_k^2, \delta\theta^2) \quad (32)$$

Hence, the NE-control scheme (13), or equivalently (16), is a first-order approximation to the gradient-based law (32). □

The drawback of NE control is that the higher-order terms are neglected as a comparison of (13) and (32) shows it. The advantage is that the experimental gradient estimation needed in (20) or (21) can be circumvented, as shown in (16) for state feedback and in the next section for output feedback.

3 Input Update via Output Feedback

NE control allows computing, to a first-order approximation, optimal input update from input and state measurements, i.e. without having to evaluate the cost sensitivities experimentally. NE control is usually based on full-state feedback. Here, an extension that accommodates output feedback will be introduced.

3.1 Output Feedback

Consider the measurement equation

$$y = M(x), \tag{33}$$

where y is the q-dimensional output vector, and its linearized form at the k^{th} iteration:

$$\delta y_k = M_x \delta x_k \tag{34}$$

Equations (8) and (34) allow writing $\delta\theta$ in terms of the output variations δy_k instead of the state variations δx_k :

$$\delta\theta = \mathcal{M}^{\dagger}(\delta y_k + M_x F_x^{-1} F_u \delta u_k) \tag{35}$$

with

$$\mathcal{M} = -M_x F_x^{-1} F_\theta \tag{36}$$

Using (35) in (13), the input update δu_{k+1} can be expressed in terms of the input and output measurements:

$$\delta u_{k+1} = K^y \delta y_k + K^u \delta u_k \tag{37}$$

with the gains

$$K^y = K\mathcal{M}^{\dagger} \tag{38}$$

$$K^u = K^y M_x F_x^{-1} F_u \tag{39}$$

Note that setting $y = x$ leads to state feedback. In this case, $M_x = I$ implying $\mathcal{M}^{\dagger} = -F_\theta^{\dagger} F_x$ and $\mathcal{M}^{\dagger} M_x F_x^{-1} F_u = -F_\theta^{\dagger} F_u$, i.e. equations (38) and (39) reduce to (17) and (18), respectively.

Note also that the NE control allows estimating the cost gradient via input and output measurements. Indeed, a comparison of (21) and (37) gives:

$$\left.\frac{d\phi}{du}\right|_{u_k} = -\left(\frac{d^2\phi}{du^2}\right)_{nom} [K^y \delta y_k + (K^u - I)\delta u_k] \tag{40}$$

3.2 Convergence Analysis of the Output Feedback Scheme

It will be shown next that, in a noise-free case and under certain conditions, the proposed output feedback update scheme converges in at most two iterations.

Theorem 1. *Under the hypothesis $\mathcal{M}^\dagger \mathcal{M} = I$, the output feedback law (37) converges locally to the optimum in at most two iterations.*

Proof. Following the disturbance $\delta\theta$ and using the update law (37), the variations δx_{k+1} can be computed from (11) and (34) as:

$$\begin{aligned}
\delta x_{k+1} &= -F_x^{-1} F_\theta \delta\theta - F_x^{-1} F_u \delta u_{k+1} \\
&= -F_x^{-1} F_\theta \delta\theta - F_x^{-1} F_u K^y M_x \delta x_k \\
&\quad -F_x^{-1} F_u K^u \delta u_k
\end{aligned}$$

The evolution of δx and δu can can be written as the following discrete-time dynamic system:

$$\begin{bmatrix} \delta x_{k+1} \\ \delta u_{k+1} \end{bmatrix} = \Phi \begin{bmatrix} \delta x_k \\ \delta u_k \end{bmatrix} + \Gamma \delta\theta \tag{41}$$

with

$$\Phi = \begin{bmatrix} -F_x^{-1} F_u K^y M_x & -F_x^{-1} F_u K^u \\ K^y M_x & K^u \end{bmatrix} \tag{42}$$

$$\Gamma = -\begin{bmatrix} F_x^{-1} F_\theta \\ 0 \end{bmatrix} \tag{43}$$

Note that Φ can be written as $\Phi = AB$ with

$$A = \begin{bmatrix} -F_x^{-1} F_u \\ I \end{bmatrix} \quad (44)$$

$$B = K^y M_x \begin{bmatrix} I & F_x^{-1} F_u \end{bmatrix} \quad (45)$$

Note also that $BA = 0$. Hence, since $\Phi^2 = ABAB = 0$, the system dynamics expressed over two time steps read:

$$\begin{bmatrix} \delta x_{k+2} \\ \delta u_{k+2} \end{bmatrix} = \Phi^2 \begin{bmatrix} \delta x_k \\ \delta u_k \end{bmatrix} + \Phi \Gamma \delta\theta + \Gamma \delta\theta = (I + \Phi)\Gamma\delta\theta \quad (46)$$

i.e. the update scheme converges locally within 2 iterations to:

$$\begin{bmatrix} \delta x^* \\ \delta u^* \end{bmatrix} = (I + \Phi)\Gamma\delta\theta \quad (47)$$

If $\mathcal{M}^\dagger \mathcal{M} = I$, it can be verified that

$$\begin{bmatrix} \delta x^* \\ \delta u^* \end{bmatrix} = (I + \Phi)\Gamma\delta\theta = \begin{bmatrix} -F_x^{-1}(F_u K + F_\theta) \\ K \end{bmatrix} \delta\theta \quad (48)$$

For a given $\delta\theta$, using the converged values $\delta x_{k+1} = -F_x^{-1}(F_u K + F_\theta)\delta\theta$ and $\delta u_{k+1} = K\delta\theta$ in (9)-(11) provides a system of redundant equation for $\delta\lambda$. It can be verified that this system has a solution, i.e. (47) is the solution of (9)-(11) for a given $\delta\theta$.

□

Remarks

1. NE control relies on the implicit estimation of the parametric disturbance $\delta\theta$ via equation (35). The feasibility of this estimation is crucial and is expressed through the condition $\mathcal{M}^\dagger \mathcal{M} = I$ in Theorem 1. This condition implies that the number of uncertain parameters

should not exceed the number of available measurements. Moreover, it is necessary that the uncertain parameters affect the measurements via the system F. If an uncertain parameter appears only in the cost function ϕ, the condition $\mathcal{M}^{\dagger}\mathcal{M} = I$ is not satisfied.

2. The convergence of measurement-based optimization schemes in the presence of time-varying parametric variations is usually based on a time-scale separation between the time variations of the parameters and the dynamics of the update scheme, i.e. the parameters must not vary significantly during the time span required for the update scheme to converge. Hence, the parameters can vary only very slowly when the number of iterations required for convergence is large, as this is usually the case with the perturbation approach. On the other hand, the proposed NE feedback converges within a few iterations, i.e. it takes much less time for convergence than the perturbation-based approach. Hence, the proposed NE feedback is better suited for handling time-varying parameters.

3.3 Dealing with Measurement Noise

In the following, the effect of noise on the performance of the NE output feedback scheme is analyzed. With the measurement noise n, the measurement equation reads:

$$y = M(x) + n \tag{49}$$

The goal is to compute the difference in cost improvement obtained with the NE output feedback scheme and the use of the nominal inputs:

$$\Delta\phi_{k+1} = \delta\phi_{k+1}^{NE} - \delta\phi_{nom} \tag{50}$$

The cost variation $\delta\phi$ indicates the difference in cost between the perturbed process with a given input update and the optimal unperturbed process. The cost variations are taken up to the second-order terms.

For the NE output feedback scheme to be useful, $\Delta\phi_{k+1}$ should be negative. This difference is a quadratic function of the parameter variations and of the noise. It is computed in the following proposition.

Proposition 2. *In the presence of the zero-mean measurement noise n and the parametric disturbance $\delta\theta$, the cost improvement by using the NE output feedback scheme compared to using the nominal optimal inputs is given by*

$$E(\Delta\phi) \quad = \quad \frac{1}{2}E(n^T\Theta_N n) + \frac{1}{2}\delta\theta^T\Theta_D\delta\theta \tag{51}$$

where $E(z)$ represents the mathematical expectation of the random variable z.

Proof. We will first compute $\delta\phi_{k+1}^{NE}$ and $\delta\phi_{nom}$ and then the difference $\Delta\phi_{k+1}$

- *Computation of $\delta\phi_{k+1}^{NE}$.* In the presence of measurement noise, the first-order variations δx and δu in (41) become:

$$\begin{bmatrix} \delta x_{k+1} \\ \delta u_{k+1} \end{bmatrix} \quad = \quad \Phi \begin{bmatrix} \delta x_k \\ \delta u_k \end{bmatrix} + \Gamma\delta\theta + AK^y n_k \tag{52}$$

For $k \geq 2$, i.e. once the noise-free system would have converged to the new optimum δx^* and δu^*, the state and input deviations of the noisy system are driven by noise only:

$$\begin{bmatrix} \delta x_{k+1} \\ \delta u_{k+1} \end{bmatrix} \quad = \quad \begin{bmatrix} \delta x^* \\ \delta u^* \end{bmatrix} + AK^y n_k \tag{53}$$

Using the notation

$$\nabla^2\phi = \begin{bmatrix} \phi_{xx} & \phi_{xu} & \phi_{x\theta} \\ \phi_{ux} & \phi_{uu} & \phi_{u\theta} \\ \phi_{\theta x} & \phi_{\theta u} & \phi_{\theta\theta} \end{bmatrix}, \tag{54}$$

and (52), the cost variation resulting from the variations δx_{k+1}, δu_{k+1} and $\delta\theta$ is given by:

$$\begin{aligned} \delta\phi_{k+1}^{NE} &= \begin{bmatrix} \phi_x & \phi_u & \phi_\theta \end{bmatrix} \begin{bmatrix} \delta x^* \\ \delta u^* \\ \delta\theta \end{bmatrix} + \begin{bmatrix} \phi_x & \phi_u \end{bmatrix} AK^y n_k \\ &+ \frac{1}{2} n_k^T \begin{bmatrix} (AK^y)^T & 0 \end{bmatrix} \nabla^2\phi \begin{bmatrix} AK^y \\ 0 \end{bmatrix} n_k \\ &+ \frac{1}{2} \begin{bmatrix} (\delta x^*)^T & (\delta u^*)^T & \delta\theta^T \end{bmatrix} \nabla^2\phi \begin{bmatrix} \delta x^* \\ \delta u^* \\ \delta\theta \end{bmatrix} \\ &+ \begin{bmatrix} (\delta x^*)^T & (\delta u^*)^T & \delta\theta^T \end{bmatrix} \nabla^2\phi \begin{bmatrix} AK^y \\ 0 \end{bmatrix} n_k \end{aligned} \tag{55}$$

Note that (55) can be simplified considering that

$$\begin{bmatrix} \phi_x & \phi_u \end{bmatrix} A = 0 \tag{56}$$

from the optimality conditions (7).

- *Computation of $\delta\phi_{nom}$.* The state variations obtained upon using the nominal inputs on the perturbed system is $\delta x_{nom} = -F_x^{-1} F_\theta \delta\theta$, and thus the nominal variation vector reads:

$$\begin{bmatrix} \delta x_{nom} \\ 0 \end{bmatrix} = \Gamma\delta\theta \tag{57}$$

Using (47) and (57), the optimal variation vector can be written as:

$$\begin{bmatrix} \delta x^* \\ \delta u^* \end{bmatrix} = \begin{bmatrix} \delta x_{nom} \\ 0 \end{bmatrix} + \Phi\Gamma\delta\theta \tag{58}$$

Furthermore, the cost variation obtained upon using the nominal inputs on the perturbed system, is given by:

$$\begin{aligned} \delta\phi_{nom} &= \begin{bmatrix} \phi_x & \phi_u & \phi_\theta \end{bmatrix} \begin{bmatrix} \delta x_{nom} \\ 0 \\ \delta\theta \end{bmatrix} \\ &+ \frac{1}{2} \begin{bmatrix} \delta x_{nom} & 0 & \delta\theta \end{bmatrix} \nabla^2\phi \begin{bmatrix} \delta x_{nom} \\ 0 \\ \delta\theta \end{bmatrix} \end{aligned} \tag{59}$$

- *Computation of* $\Delta\phi_{k+1}$.

It follows from $\phi = AB$ and (56) that

$$\begin{bmatrix} \phi_x & \phi_u \end{bmatrix} \phi = 0 \tag{60}$$

and, with (57) and (58):

$$\begin{bmatrix} \phi_x & \phi_u & \phi_\theta \end{bmatrix} \begin{bmatrix} \delta x^* \\ \delta u^* \\ \delta\theta \end{bmatrix} = \begin{bmatrix} \phi_x & \phi_u & \phi_\theta \end{bmatrix} \begin{bmatrix} \delta x_{nom} \\ 0 \\ \delta\theta \end{bmatrix} \tag{61}$$

Hence, the improvement in cost by using NE feedback compared to the nominal input values is:

$$\begin{aligned}\Delta\phi_{k+1} &= \frac{1}{2}n_k^T\Lambda^T\nabla^2\phi\Lambda n_k \\ &+ \frac{1}{2}\delta\theta^T(\Psi+\Delta)^T\nabla^2\phi(\Psi+\Delta)\delta\theta \\ &- \delta\theta^T\Psi^T\nabla^2\phi\Psi\delta\theta \qquad (62) \\ &+ \delta\theta^T(\Psi+\Delta)^T\nabla^2\phi\Lambda n_k \qquad (63)\end{aligned}$$

with

$$\Lambda = \begin{bmatrix} AK^y \\ 0 \end{bmatrix} \qquad (64)$$

$$\Delta = \begin{bmatrix} \Phi\Gamma \\ 0 \end{bmatrix} \qquad (65)$$

$$\Psi = \begin{bmatrix} \Gamma \\ I \end{bmatrix} \qquad (66)$$

Upon defining the parameters:

$$\begin{aligned}\Theta_D &= \Psi^T\nabla^2\phi\Delta + \Delta^T\nabla^2\phi\Psi + \Delta^T\nabla^2\phi\Delta \\ \Theta_N &= \Lambda^T\nabla^2\phi\Lambda \qquad (67) \\ \Theta_I &= (\Psi+\Delta)^T\nabla^2\phi\Lambda \qquad (68)\end{aligned}$$

the variation $\Delta\phi_{k+1}$ reads:

$$\Delta\phi_{k+1} = \frac{1}{2}n_k^T\Theta_N n_k + \frac{1}{2}\delta\theta^T\Theta_D\delta\theta + \delta\theta^T\Theta_I n_k \qquad (69)$$

For a zero-mean noise n, taking the mathematical expectation of the cost variation $E(\Delta\phi)$ in (69) gives (51).

□

Remarks

1. Input update is useful only if the cost variation due to measurement noise is small compared to the cost variation due to parametric disturbances. The parameter ρ describes the 'noise-to-signal' ratio of the cost variation (provided $\delta\theta$ and n are independent stochastic variables, and n is a zero-mean noise) :

$$\rho = \| \frac{E(n_k^T \Theta_N n_k)}{E(\delta\theta^T \Theta_D \delta\theta)} \| \tag{70}$$

   For $\rho << 1$, the noise level can be considered negligible for the efficiency of the NE-feedback scheme. If ρ is close to or greater than 1, the measurement noise renders the input update irrelevant. The parameter ρ can be understood as follows: If the noise-free NE scheme converges to a point close to the true optimum ϕ^*, i.e. $\phi^{NE}_{Noise-free} \approx \phi^*$, then

$$\phi^* - \phi_{nom} \approx \phi^{NE}_{Noise-free} - \phi_{nom} \approx \frac{1}{2}\delta\theta^T \Theta_D \delta\theta \tag{71}$$

   Moreover

$$E(\Delta\phi^{NE}) \approx E(\phi^{NE} - \phi_{nom}) \tag{72}$$

   Hence, the parameter ρ allows assessing performance of NE feedback compared to the maximally possible improvement:

$$\frac{E(\phi^{NE} - \phi_{nom})}{E(\phi^* - \phi_{nom})} \approx \frac{E(\Delta\phi^{NE})}{E(\phi^* - \phi_{nom})} \approx 1 - \rho \tag{73}$$

   Note that all the approximations used in (71)-(73) are first-order approximations.

2. If the distributions of the parametric disturbance $\delta\theta$ and measurement noise n are known, expression (70) can be computed off-line, and the relevance of the input update scheme evaluated before the algorithm is actually implemented.

3. The matrix Θ_D quantifies the influence of the various parametric disturbances on the loss of optimality, while the matrix Θ_N quantifies the influence of measurement noise on the efficiency of the NE-feedback scheme.

4 Optimization of a Continuous Chemical Reactor

4.1 System Description

An isothermal continuous chemical reactor with the two reactions $A + B \rightarrow C$ and $2\,B \rightarrow D$ is considered. There are two feeds with the flow rates F_A and F_B of concentrations $c_{A_{in}}$ and $c_{B_{in}}$, respectively. The optimization objective is to compute the flow rates F_A and F_B that maximize the amount of C produced, $(F_A + F_B)c_C$, while also considering the selectivity factor $\frac{(F_A+F_B)c_C}{F_A c_{A_{in}}}$ and the weighted norm of the input flow rates $\frac{1}{2}\begin{bmatrix} F_A & F_B \end{bmatrix} R \begin{bmatrix} F_A & F_B \end{bmatrix}^T$. The optimization problem can be formulated mathematically as:

$$\max_{F_A, F_B} \phi = (F_A + F_B)c_C \frac{(F_A + F_B)c_C}{F_A c_{A_{in}}} - \frac{1}{2}\begin{bmatrix} F_A & F_B \end{bmatrix} R \begin{bmatrix} F_A \\ F_B \end{bmatrix} \tag{74}$$

$$s.t. \quad F_A c_{A_{in}} - (F_A + F_B)c_A - k_1 c_A c_B V = 0 \tag{75}$$

$$F_B c_{B_{in}} - (F_A + F_B)c_B - k_1 c_A c_B V - 2k_2 c_B^2 V = 0 \tag{76}$$

$$-(F_A + F_B)c_C + k_1 c_A c_B V = 0 \tag{77}$$

where c_X is the concentration of species X, and k_1, k_2 the rate constants. The nominal model parameters and operating conditions are given in Table 1.

The parameter k_1 and k_2 are considered uncertain in the ranges [0.3 - 3] $\frac{\mathrm{l}}{\mathrm{mol\,h}}$ and [10^{-3} - 10^{-1}]

k_1	0.65	$\frac{\text{l}}{\text{mol h}}$		$c_{B_{in}}$	1.5	$\frac{\text{mol}}{\text{l}}$
k_2	0.014	$\frac{\text{l}}{\text{mol h}}$		V	500	l
$c_{A_{in}}$	2	$\frac{\text{mol}}{\text{l}}$				

Table 1: Model parameters and operating conditions.

$\frac{\text{l}}{\text{mol h}}$, respectively. Note that the range for k_2 is rather large. The input weighting matrix R is chosen as the identity matrix $I_{2\times 2}$.

4.2 Output NE feedback - Sufficient Measurements

If the number of measurements is sufficient to reconstruct the uncertain parameters according to (35), condition $\mathcal{M}^\dagger\mathcal{M} = I$ is fulfilled. Consider the case with two output measurements, $y = \begin{bmatrix} c_B & c_C \end{bmatrix}^T$, which allows taking into account the uncertainty in both k_1 and k_2. The feedback gains obtained for the nominal operating point corresponding to $k_1 = 0.65$ and $k_2 = 0.014$ are:

$$
\begin{aligned}
K^y &= \begin{bmatrix} 0.4207 & -1.7303 \\ 0.0690 & -2.4185 \end{bmatrix} \\
K^u &= \begin{bmatrix} 0.7734 & -0.6366 \\ 0.8484 & -0.7027 \end{bmatrix}
\end{aligned}
$$

Figure 1 compares, for different values of the uncertain parameters k_1 and k_2, the cost obtained with NE feedback, the cost obtained using the nominal inputs, and the true optimal cost. The true optimum has been computed numerically assuming perfect knowledge of the parameters k_1 and k_2. The cost obtained with the NE feedback differs from the true optimum cost by less than 0.025% for all values of k_1 and k_2, although the controller is computed for $k_1 = 0.65$ and $k_2 = 0.014$ and does not know the value of k_1 and k_2 used in the simulation. Applying the nominal input to the perturbed system leads to sub-optimality (up to 8.5% loss in cost). As expected, the loss in cost increases with the size of the parametric uncertainty.

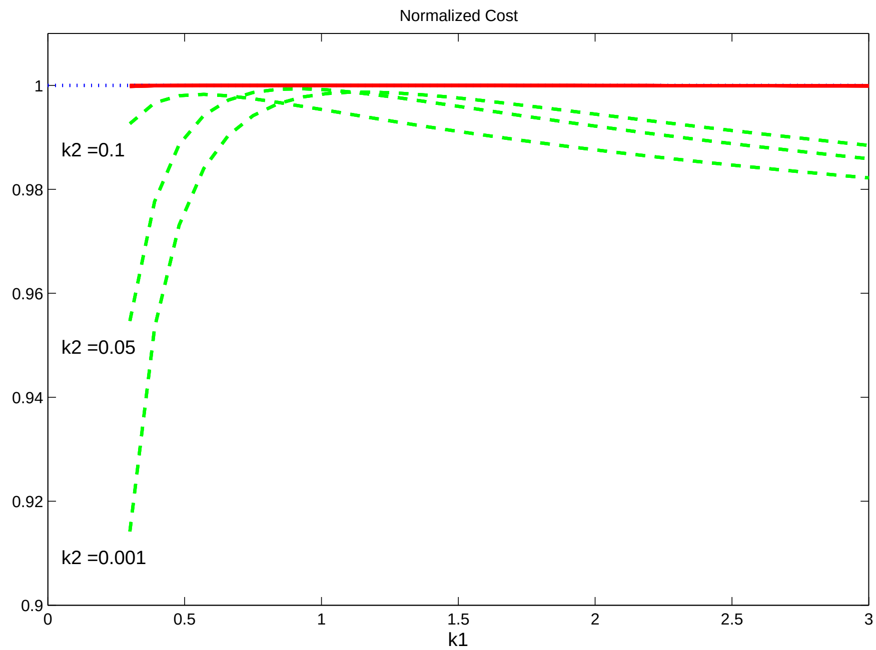

Figure 1: Comparison of the costs obtained with NE feedback, direct application of the nominal inputs, and true optimal cost, for different values of the uncertain parameters k_1 and k_2. States c_B and c_C are measured. The various costs are normalized with respect to the true optimal cost so that the optimum is always 1. *Dotted line*: true optimal cost; *dashed line*: direct application of the nominal inputs; *solid line*: NE control after 3 iterations. Note that the cost obtained using NE feedback cannot be distinguished from the true optimum.

4.3 Output NE feedback - Insufficient measurements

Consider the case with a single measurement, $y = c_C$. To satisfy the condition $\mathcal{M}^\dagger \mathcal{M} = I$, only one uncertain parameter can be considered for controller design. The cost being more sensitive to uncertainties in k_1 than in k_2, uncertainties in k_2, although present in the system, will be ignored in designing the control law.

The feedback gains obtained for the nominal operating point corresponding to $k_1 = 0.65$ and $k_2 = 0.014$ are:

$$
\begin{aligned}
K^y &= \begin{bmatrix} -1.9244 \\ -2.4503 \end{bmatrix} \\
K^u &= \begin{bmatrix} 0.6505 & -0.5394 \\ 0.8282 & -0.6867 \end{bmatrix}
\end{aligned}
$$

Figure 2 compares, for different values of the uncertain parameters k_1 and k_2, the cost obtained with NE feedback, the cost obtained using the nominal inputs, and the true optimal cost. The cost obtained using NE feedback is close to the true optimal cost although the number of outputs is insufficient to guarantee convergence (the difference is less than 0.62% for all values of k_1 and k_2).

4.4 Output NE feedback with Measurement Noise

Consider the case with two outputs as in Section 4.2, but with additive measurement noise, i.e.

$$y = M(x) + n \tag{78}$$

The feedback gains obtained for the nominal operating point corresponding to $k_1 = 0.65$ and $k_2 = 0.014$ are the same as in Section 4.2 The matrices Θ_D and Θ_N are:

$$
\begin{aligned}
\Theta_D &= \begin{bmatrix} -0.0166 & 0.1812 \\ 0.1812 & -2.1059 \end{bmatrix} \\
\Theta_N &= \begin{bmatrix} -0.1441 & 0.0119 \\ 0.0119 & 10.2615 \end{bmatrix}
\end{aligned}
$$

Three cases are presented next:

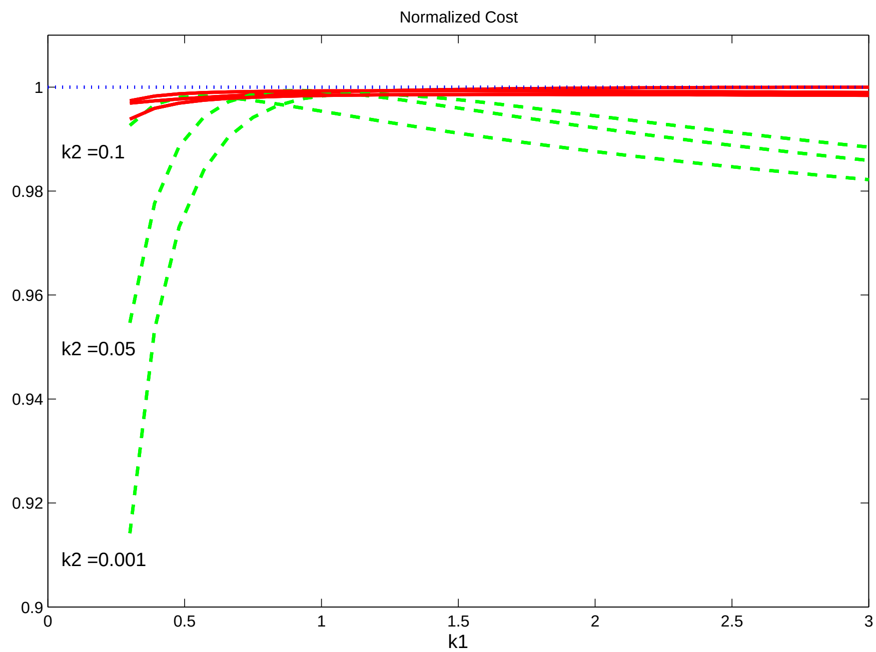

Figure 2: Comparison of the costs obtained with NE feedback, direct application of the nominal inputs, and true optimal cost, for different values of the uncertain parameters k_1 and k_2. State c_C is measured. The uncertainty in k_2 is ignored when designing the NE controller. The various costs are normalized with respect to the true optimal cost so that the optimum is always 1. *Dotted line*: true optimal cost; *dashed line*: direct application of the nominal inputs; *solid line*: NE control after 3 iterations. The cost obtained using NE control differs slightly from the true optimum.

- Figure 3 compares the performance of three scenarios for large uncertainties in k_1 and k_2, and zero-mean gaussian measurement noise of 2% standard deviation. The noise-to-signal ratio is $\rho = 1.8 \cdot 10^{-4}$. The performance of NE control is computed according to (73) by averaging 1000 realizations of ϕ^{NE}, with the result $\frac{E(\phi^{NE}-\phi_{nom})}{E(\phi^{*}-\phi_{nom})} = 0.996$. The proposed scheme works well in the present case.

- For the same amount of measurement noise, Figure 4 compares the performance of three scenarios for 20% deviation in k_1 ($k_1 = 0.78$; $k_2 = k_2^{nom}$). The noise-to-signal ratio is $\rho = 0.04$, i.e. the level of measurement noise is now small but not negligible with regard to the level of parametric uncertainty. The performance of NE control is computed according to (73), with the result $\frac{E(\phi^{NE}-\phi_{nom})}{E(\phi^{*}-\phi_{nom})} = 0.972$. The proposed scheme works well in the present case, yet it shows a slight loss of performance when compared to the previous case.

- Figure 5 compares the performance of three scenarios for 20% deviation in k_1 ($k_1 = 0.78$; $k_2 = k_2^{nom}$) and zero-mean gaussian measurement noise of 10% standard deviation. The noise-to-signal ratio is $\rho = 1.01$, i.e. the level of measurement noise is now critical with regard to the level of parametric uncertainty. The performance of NE control is computed according to (73), with the result $\frac{E(\phi^{NE}-\phi_{nom})}{E(\phi^{*}-\phi_{nom})} = 0.291$, i.e. the performance of the proposed scheme is poor in the present case.

It can be seen from these examples that, for a fixed amount of parametric uncertainty, the efficiency of the proposed scheme decreases with the amount of measurement noise. Conversely, for a fixed amount of measurement noise, the relative efficiency of the proposed scheme, as defined by (73), *increases* with parametric uncertainty.

5 Conclusion

The optimization of static processes has been addressed for the case of model mismatch but availability of measurements. NE output feedback is used as an alternative to the classical perturbation approach. It relies on more information (output instead of simply the cost function) and thus requires far fewer iterations to reach the optimum. The symbolic and numerical computations required to compute the feedback are rather straightforward and can be performed off-line.

A predictor of the performance of the NE scheme in the presence of measurement noise is proposed. This predictor can be computed off-line if the stochastic properties of the uncertain parameters and

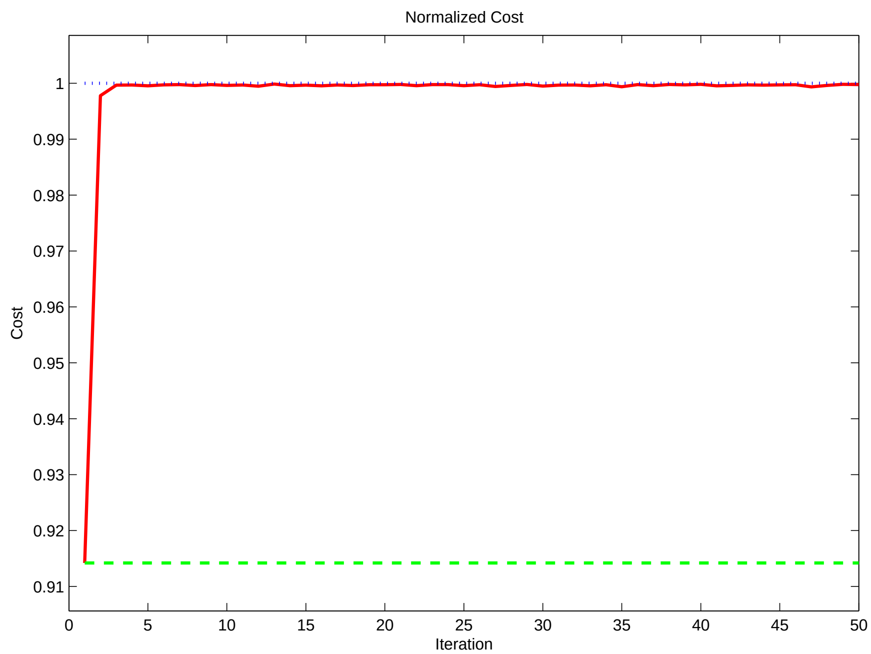

Figure 3: Comparison of the costs obtained with NE feedback, direct application of the nominal inputs, and true optimal cost, for $k_1 = 0.3$ and $k_2 = 0.001$, and 2% of zero-mean gaussian measurement noise. The various costs are normalized with respect to the true optimal cost so that the optimum is always 1. *Dotted line*: true optimal cost; *dashed line*: direct application of nominal inputs; *solid line*: NE control after 3 iterations.

the measurement noise can be assessed. Though the accuracy of the predictor is limited by first-order approximations, it is consistent with the behavior observed in simulations.

The main drawback of the NE approach is that it only represents a first-order approximation and, furthermore, it does not necessarily converge to the true optimum in the presence of uncertainty. Note that the proof of convergence has been obtained for the nominal case. Although the resulting non-optimality is often minor (like in the reactor example that has been considered), it can be unacceptable in certain scenarios.

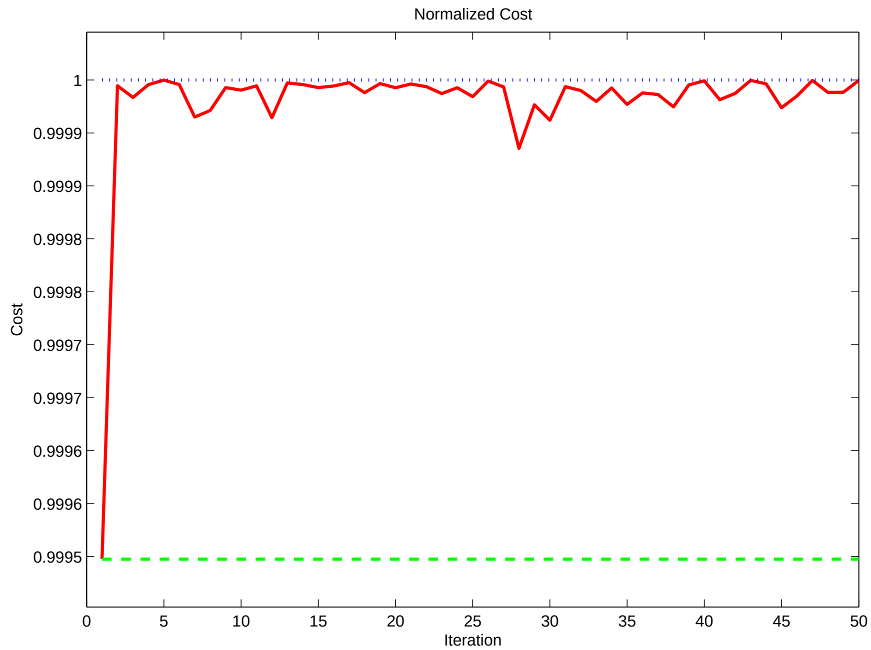

Figure 4: Comparison of the costs obtained with NE feedback, direct application of the nominal inputs, and true optimal cost, for $k_1 = 0.78$ (i.e. 20% deviation), $k_2 = k_2^{nom}$ and 2% zero-mean gaussian measurement noise. The various costs are normalized with respect to the true optimal cost so that the optimum is always 1. *Dotted line*: true optimal cost; *dashed line*: direct application of nominal inputs; *solid line*: NE control after 3 iterations.

References

Abel, O. and W. Marquardt (1998). A model predictive control scheme for safe and optimal operation of exothermic semi-batch reactors. In: *IFAC DYCOPS-5*. Corfu, Greece. pp. 761–766.

Box, G. E. P. and N. R. Draper (1987). *Empirical Model-building and Response Surfaces*. John Wiley, New York.

Bryson, A. E. and Y. C. Ho (1975). *Applied Optimal Control*. Hemisphere, Washington DC.

Eaton, J. W. and J. B. Rawlings (1990). Feedback control of nonlinear processes using on-line optimization techniques. *Comp. Chem. Eng.* **14**, 469–479.

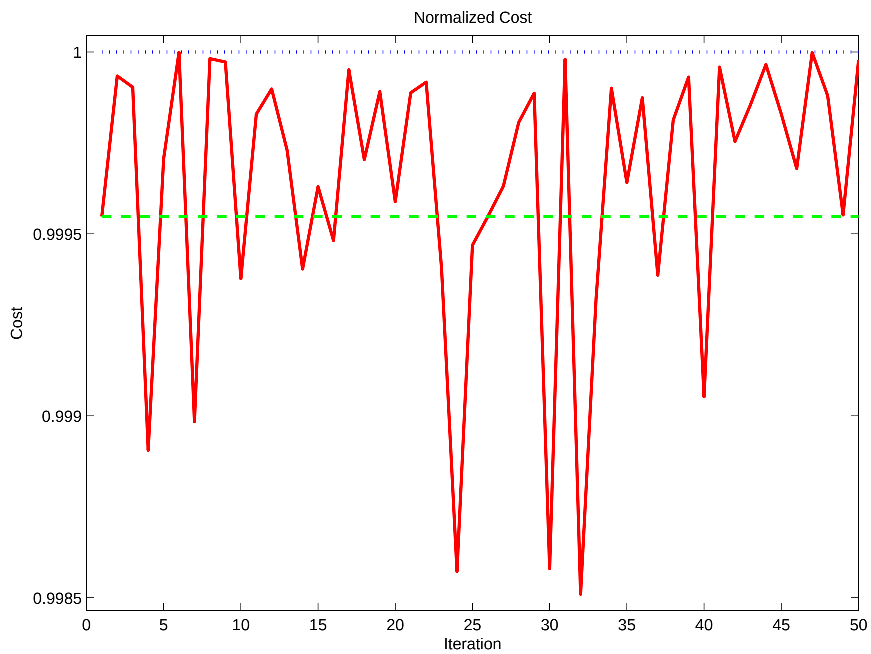

Figure 5: Comparison of the costs obtained with NE feedback, direct application of the nominal inputs, and true optimal cost, for $k_1 = 0.78$ (i.e. 20% deviation), $k_2 = k_2^{nom}$ and 10% zero-mean gaussian measurement noise. The various costs are normalized with respect to the true optimal cost so that the optimum is always 1. *Dotted line*: true optimal cost; *dashed line*: direct application of nominal inputs; *solid line*: NE control after 3 iterations.

François, G., B. Srinivasan and D. Bonvin (2005). Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty. *J. Process Contr.* **15**(6), 701–712.

Kadam, J. and W. Marquardt (2004). Sensitivity-based solution updates in closed-loop dynamic optimization. In: *IFAC DYCOPS-7*. Boston, Massachusetts.

Kadam, J. V., W. Marquardt, M. Schlegel, O. H. Bosgra T. Backx, P.-J. Brouwer, G. Dünnebier, D. van Hessem, A. Tiagounov and S. de Wolf (2003). Towards integrated dynamic real-time optimization and control of industrial processes. In: *FOCAPO 2003, Fourth International*

Conference on Foundations of Computer-Aided Process Operations. Coral Springs, Florida.

Kristic, M. and H-H. Wang (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica* **36**, 595–601.

Marlin, T. and A. N. Hrymak (1996). Real-time operations optimization of continuous processes. In: *Chemical Process Control-V Conference*. Tahoe City, Nevada.

Morosanov, I. I. (1957). Method of extremum control. *Automatic and Remote Control* **18**, 1077–1092.

Ostrovskii, I. I. (1957). Extremum regulation. *Automatic and Remote Control* **18**, 900–907.

Pervozvanskii, A. A. (1960). Continuous extremum control system in the presence of random noise. *Automatic and Remote Control* **21**, 673–677.

Ruppen, D., C. Benthack and D. Bonvin (1995). Optimization of batch reactor operation under parametric uncertainty - Computational aspects. *J. Process Contr.* **5**(4), 235–240.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure. *J. Process Contr.* **10**, 487–507.

Srinivasan, B., D. Bonvin, E. Visser and S. Palanki (2003). Dynamic optimization of batch processes: II. Role of measurements in handling uncertainty. *Comp. Chem. Eng.* **44**, 27–44.

Zang, T., M. Guay and D. Dochain (2001). Adaptive extremum seeking control of continuous stirred tank bioreactors. *AIChE J.* **40**(2), 10–20.