Self-Optimizing Control of a Continuous-Flow Pharmaceutical Manufacturing Plant

David Pérez Piñeiro * **Anastasia Nikolakopoulou** **
Johannes Jäschke * **Richard D. Braatz** **

* *Norwegian University of Science and Technology, NO-7491 Trondheim (e-mail: david.p.pineiro@ntnu.no, johannes.jaschke@ntnu.no).*
** *Massachusetts Institute of Technology, Cambridge, MA 02139, USA (e-mail: anikol@mit.edu, braatz@mit.edu).*

Abstract: This article considers the real-time optimization under uncertainty of a compact reconfigurable system for on-demand continuous-flow pharmaceutical manufacturing. Self-optimizing control is employed, which optimizes operation in the presence of uncertainty by controlling a carefully chosen combination of measurements to a constant setpoint. The method is applied to a simulated plant based on the physical process. The closed-loop simulations indicate that this simple policy is able to maintain the process operation close to optimality despite disturbances, sensor noise, and parametric model uncertainty.

Keywords: Continuous manufacturing, pharmaceuticals, real-time optimization, plantwide control, self-optimizing control.

1. INTRODUCTION

Pharmaceuticals have been traditionally manufactured using batch processes. Challenges facing the pharmaceutical industry such as drug shortages often trace back to limitations of batch manufacturing, which have motivated efforts to develop continuous manufacturing processes (Myerson et al., 2015). Compact reconfigurable systems for on-demand continuous-flow pharmaceutical manufacturing (Adamo et al., 2016) have been designed that include all of the molecular synthesis and separation steps for drug manufacturing. Such systems can respond to sudden changes in demand, such as in cases of epidemics or pandemics, by manufacturing drugs on-demand at high volumetric efficiencies (Adamo et al., 2016).

This article considers the design of a self-optimizing control system for the modular drug manufacturing platform of Adamo et al. (2016) for optimizing plant operation in the presence of different forms of uncertainty. In past work, a model predictive control-based plantwide control system (Nikolakopoulou et al., 2019) was designed for controlling overall yield and production rate while satisfying operational constraints.

In pharmaceutical processes, the effects of model uncertainty (both in the structure and the parameters), sensor noise, and process disturbances such as exogenous changes in temperatures and concentrations must be considered. An important point of departure from deterministic optimal control problems is that our goal is not to optimize over control sequences, but over control policies or feedback control laws. A control policy is a function that determines a feasible control action given the state of the system at a point in time. There are two fundamental strategies for designing policies (Powell, 2019):

- Policy search, where an objective function is used to search within a family of functions to find the function that works best.
- Lookahead approximations, where the policy is constructed now by approximating the impact of a control action on the future.

Either approach can produce optimal policies, although this is rare. The reason is computational; a stochastic dynamic program can rarely be solved to optimality. However, these two strategies are the basis for 4 universal approximations (policy function, cost function, value function, and direct lookahead) that cover all of the approaches reported in the literature (Powell, 2019). Economic model predictive control is a policy based on direct lookahead approximation that is receiving much research attention. Alternatively, this article explores the potential of an approach to a policy search known as "self-optimizing control" which, despite its ease of implementation, has never been applied to pharmaceutical manufacturing.

This paper is organized as follows. Section 2 presents self-optimizing control and the algorithms. Section 3 introduces the case study, which is the continuous-flow synthesis of atropine. Section 4 contains the results and discussion. The article concludes in Section 5.

2. SELF-OPTIMIZING CONTROL

Self-optimizing control is a policy search approach which optimizes the process operation in the presence of uncertainty by controlling a carefully chosen combination of

measurements to a constant setpoint (Morari et al., 1980; Skogestad, 2000). The control architecture when applying self-optimizing control is illustrated in Fig. 1.

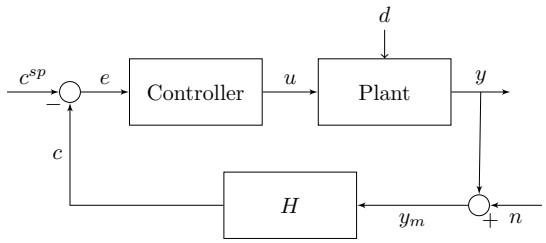

Fig. 1. Block diagram of a self-optimizing control architecture. The variables in the diagram include disturbances d, true measurements y, sensor noise n, measurements corrupted by noise y_m, combination matrix H, controlled variables c, setpoints c^{sp}, error e, and manipulated variables u.

Given this structure, the problem of optimizing over policies is simplified to the search for the best combination matrix H and vector of setpoints c^{sp}. These two parameters define the self-optimizing control block. The design of the feedback controller (e.g., decentralized PID controllers or MPC) is considered separately. The search for the parameters H and c^{sp} can be combined into one problem by manipulating notation. Defining the augmented combination matrix by

$$\tilde{H} = [-c^{sp} \;\; H] \tag{1}$$

and the augmented vector of measurements by

$$\tilde{y}^\top = \begin{bmatrix} 1 & y^\top \end{bmatrix}, \tag{2}$$

the problem of controlling $c = Hy$ to an arbitrary setpoint c^{sp} is equivalent to controlling $\tilde{c} = \tilde{H}\tilde{y}$ to a zero setpoint. Therefore, an equivalent formulation of the problem is to find the best augmented combination matrix $\tilde{H}$. This problem is considered in the remainder of this article and, to simplify notation, the augmented variables are written without the tildes.

Several model- and data-based methods have been developed to search for the best combination matrix H efficiently. A comprehensive survey on self-optimizing control methods is available (Jäschke et al., 2017).

2.1 A global self-optimizing control algorithm

Consider two different forms of uncertainty: sensor noise n and disturbances d. Other types of uncertainty such as parametric model uncertainty can be effectively modeled as disturbances within this framework.

In self-optimizing control, a loss function is used as a performance criterion to evaluate different combination matrices H. This loss function is defined as the difference between the cost function at steady-state resulting from using a particular H and the optimal cost function at steady-state corresponding to a particular disturbance:

$$L(n, d, H) = J(n, d, H) - J^{\text{opt}}(d). \tag{3}$$

Assuming a given probability distribution for the sensor noise and disturbance realizations, the average loss can be computed as

$$L_{\text{av}}(H) = \mathbb{E}_{n,d}[L(n, d, H)]. \tag{4}$$

Therefore, the combination matrix H that minimizes the average loss can be obtained by solving

$$\min_H \;\; L_{\text{av}}(H) \tag{5a}$$
$$\text{s.t.} \;\; y = f(u, d), \tag{5b}$$
$$H(y + n) = 0, \tag{5c}$$

where the constraints are the process model and the feedback control effects, respectively. This optimization is not tractable. A convex approximation used in the literature to simplify the optimization linearizes the process model around a nominal operating point and uses a second-order Taylor expansion to approximate the loss. Then the loss can be written as an explicit function of H. The approaches based on these simplifications are collectively known as "local methods", because the policy performance is limited to a neighborhood of the nominal operating point. More recently, "global methods" that do not require the linearization of the process model around a nominal operating point have been considered. Instead, Monte Carlo simulation is used to approximate the nonlinear operating region. This article uses a global self-optimizing control algorithm proposed by Ye et al. (2015). An outline of this algorithm is:

(1) Sample the disturbance space $\mathcal{D}$ via Monte Carlo simulation, generating a sequence of N disturbance scenarios: d_i, $i = 1, \ldots, N$.
(2) For each disturbance scenario d_i, compute the corresponding optimal vector of measurements y_i^{opt}. Then, construct the intermediate matrices:

$$Y = \begin{bmatrix} (y_1^{\text{opt}})^\top \\ \vdots \\ (y_N^{\text{opt}})^\top \end{bmatrix}, \quad \tilde{Y} = \begin{bmatrix} \frac{1}{\sqrt{N}} Y \\ W_n \end{bmatrix}, \tag{6}$$

where $W_n^2 = \mathbb{E}[nn^\top]$ is the covariance matrix of Gaussian sensor noise.
(3) Evaluate $G^y = \partial y / \partial u$ and $J_{uu}^{1/2} = (\partial^2 J / \partial u^2)^{1/2}$ at a chosen reference point (e.g., at the nominal operating conditions).
(4) Formulate the convex optimization in terms of the approximated average loss $\bar{L}_{\text{av}}$:

$$\min_H \;\; \bar{L}_{\text{av}}(H) = \tfrac{1}{2} \|\tilde{Y} H^\top\|_F^2 \tag{7a}$$
$$\text{s.t.} \;\; HG^y = J_{uu}^{1/2} \tag{7b}$$

(5) The optimal H that minimizes $\bar{L}_{\text{av}}$ can be analytically computed by

$$H^\top = (\tilde{Y}^\top \tilde{Y})^{-1} G^y (G^{y\top} (\tilde{Y}^\top \tilde{Y})^{-1} G^y)^{-1} J_{uu}^{1/2} \tag{8}$$

2.2 Selecting subset of measurements

In plants with a large number of potential measurements, good performance can often be obtained by selecting only a subset of measurements. The performance does not improve linearly with the number of measurements included in the combination matrix H, but follows a Pareto front. As such, there is a point where a further performance improvement is marginal in relation with the cost of adding a new sensor (Kariwala et al., 2008).

The problem of searching over subsets of measurements is a combinatorial optimization that has been addressed by

two different approaches: tailor-made branch-and-bound algorithms (Kariwala and Cao, 2010) and reformulation as a mixed integer quadratic optimization problem (MIQP) and then applying standard MIQP solvers to find the best subset of measurements (Yelchuru and Skogestad, 2012). This article uses this second approach to reformulate the global self-optimizing control problem introduced in (7) to search over subsets of measurements.

To use a standard MIQP solver such as `CPLEX` or `Gurobi`, first reformulate the combination matrix

$$H = \begin{bmatrix} -c_1^{sp} & h_{11} & \dots & h_{1n_y} \\ \vdots & \vdots & \ddots & \vdots \\ -c_{n_u}^{sp} & h_{n_u 1} & \dots & h_{n_u n_y} \end{bmatrix} \tag{9}$$

by stacking the rows of H in a column vector,

$$h_\delta = \begin{bmatrix} -c_1^{sp} & h_{11} \cdots h_{1n_y}, & \cdots, & -c_{n_u}^{sp} & h_{n_u 1} \cdots h_{n_u n_y} \end{bmatrix}^\top \tag{10}$$

Restructuring the other matrices in a similar way results in a vectorized version of problem (7):

$$\min_{h_\delta} \quad h_\delta^\top Y_\delta h_\delta \tag{11a}$$

$$\text{s.t.} \quad {G_\delta^y}^\top h_\delta = j_\delta, \tag{11b}$$

where G_δ^y, j_δ, and Y_δ are the restructured versions of G^y, J_{uu} and Y, respectively. Problem (11) can be reformulated as a mixed integer quadratic program by adding a constraint to limit the number of measurements (number of columns in matrix H) from which the optimizer can select the best subset. These constraints can be implemented using big-M constraints. Define a vector of binary variables

$$\sigma = \begin{bmatrix} 1, \sigma_1, & \dots, & \sigma_{n_y} \end{bmatrix}, \quad \sigma_j = \{0, 1\}, \tag{12}$$

where the first element is 1 because the first column of the augmented combination matrix introduced in (1) contains the setpoints of the controlled variables, which are always selected. For the rest of the columns in the combination matrix, $\sigma_j = 1$ if measurement j is selected, and $\sigma_j = 0$ otherwise. The constraint on the binary variables can be written as

$$P\sigma = s, \tag{13}$$

where $P = \underline{1}_{1\times n_y}^\top$ is a n_y dimensional vector of ones, and s is the number of measurements we want to include in the combination matrix.

The MIQP problem for selecting the optimal combination matrix with a subset of s measurements can be written as

$$\min_{h_\delta, \sigma} \quad h_\delta^\top Y_\delta h_\delta \tag{14a}$$

$$\text{s.t.} \quad {G_\delta^y}^\top h_\delta = j_\delta \tag{14b}$$

$$P\sigma = s \tag{14c}$$

$$\sigma_1 = 1 \tag{14d}$$

$$-M\sigma_j \le h_{ij} \le M\sigma_j \quad \forall j \in \{1, \dots, n_y\}, \ \forall i \in \{1, \dots, n_u\} \tag{14e}$$

$$h_\delta \in \mathbb{R} \tag{14f}$$

$$\sigma \in \{0, 1\}. \tag{14g}$$

The vector $M \in \mathbb{R}_+^{n_u}$ of positive constants used in the big-M constraints ensures that if $\sigma_j = 0$, the corresponding column in the H matrix is also zero, and therefore the measurement j is not selected. However, selecting appropriate values for M is not straightforward. A large value implies a large upper bound on the elements of H and therefore the computational load can become very high, while a small value may result in a falsely active constraint and therefore a suboptimal solution. One strategy is to solve the problem with increasing values of M iteratively, until no changes are seen in the solution. Alternatively, it is also possible to replace big-M constraints for indicator constraints:

$$\sigma_j = 0 \Rightarrow h_{ij} = 0 \quad \forall j \in \{1, \dots, n_y\}, \ i \in \{1, \dots, n_u\}. \tag{15}$$

To select more than one subset of measurements for further validation, adding the constraint

$$(\sigma^{l-1})^\top \sigma^l \le s \quad \forall l \in \{2, \dots, m\}. \tag{16}$$

to the problem ensures that the best m subsets of measurements with least losses in increasing order are selected.

Table 1. Chemical species in the simulation.

Chemical species	Chemical formula	Notation
Tropine	$C_8H_{15}NO$	C_1
Dimethylformamide	C_3H_7NO	C_2
Phenylacetylchloride	C_8H_7ClO	C_3
Intermediate	$C_{16}H_{21}O_2NHCl$	C_4
Formaldehyde	CH_2O	C_5
Methanol	CH_3OH	C_6
Sodium hydroxide	NaOH	C_7
Water	H_2O	C_8
Atropine	$C_{17}H_{23}NO_3$	C_9
Apoatropine	$C_{17}H_{21}NO_2$	C_{10}
Tropine ester	$C_{16}H_{21}O_2N$	C_{11}
Sodium chloride	NaCl	C_{12}
Buffer	NH_4Cl	C_{13}
Toluene	C_7H_8	C_{14}

3. APPLICATION: CONTINUOUS-FLOW SYNTHESIS OF ATROPINE

This article uses a first-principles model of the continuous-flow synthesis of atropine (Nikolakopoulou et al., 2020), which is an active pharmaceutical ingredient. The process flowsheet in Fig. 2 shows the manipulated volume flowrates, reactor temperatures, and unit operations, which are three mixers (Mi), three tubular reactors (TRi), and a liquid-liquid separator (LL). Each unit operation is described by mass balances, reaction stoichiometry, and reaction kinetics. The energy balances are not modeled since the reactors can be kept at the desired temperature setpoint using regulatory level controllers. The setpoint temperature can be achieved rapidly due to the high surface-to-volume ratio of the tubular reactors.

The six feed streams to the process are:

- Tropine in dimethylformamide (DMF, q_1)
- Phenylacetylchloride (q_2)
- Formaldehyde (q_3)
- Sodium hydroxide (NaOH, q_4)
- Buffer solution (q_5)
- Organic solvent (q_6)

Each stream contains up to 14 chemical species (Table 1). Atropine (C_9) is produced according to the reaction set:

$$C_1 + C_3 \rightarrow C_4 \tag{17a}$$

$$C_4 + C_7 \rightarrow C_8 + C_{11} + C_{12} \tag{17b}$$

$$C_5 + C_{11} \rightarrow C_9 \tag{17c}$$

$$C_5 + C_{11} \rightarrow C_8 + C_{10} \tag{17d}$$

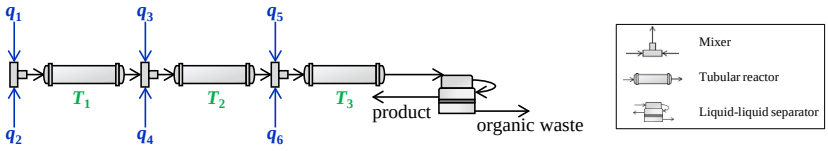

Fig. 2. Process flowsheet of the continuous-flow synthesis of atropine showing three mixers (M), three tubular reactors (TR), and one liquid-liquid separator (LL). The volume flowrates of the feed streams are q_i and the reactor temperates are indicated by T_i.

Table 2. Potential process measurements.

Index	Measurement
1-2	Concentration of C_i at the outlet of TR1, $i = \{1, 3\}$
3-10	Concentration of C_i at the outlet of TR2, $i = \{1, 3, 5, 7, 9, 10, 11, 12\}$
11-18	Concentration of C_i at the outlet of TR3, $i = \{1, 3, 5, 7, 9, 10, 11, 12\}$
19-22	Concentration of C_i in LL (aqueous phase), $i = \{1, 7, 9, 12\}$
23-26	Concentration of C_i in LL (organic phase), $i = \{3, 5, 10, 11\}$
27-30	Concentration of C_i in the feed streams q_{1-4}, $i = \{1, 3, 5, 7\}$
31-33	Volume flowrates of TR1, TR2, and TR3
34-35	Volume flowrates of aqueous and organic phases in LL
36-39	Volume flowrates of feed streams q_{1-4}
40-42	Temperatures of TR1, TR2, and TR3

Table 3. Expected values (μ) and standard deviations (σ) of disturbances.

Disturbance	Unit	μ	σ
M_{C_1}	mol/L	2	0.01μ
M_{C_7}	mol/L	4	0.01μ
T_1	K	373.15	1
T_2	K	373.15	1
T_3	K	323.15	1
k_1	mol/(mL · min)	34206	0.05μ
k_2	mol/(mL · min)	10000	0.05μ
k_3	mol/(mL · min)	24	0.05μ
k_4	mol/(mL · min)	43599	0.05μ
E_{A1}	J/mol	1000	0.05μ
E_{A2}	J/mol	100	0.05μ
E_{A3}	J/mol	1819	0.05μ
E_{A4}	J/mol	26207	0.05μ
$\log(D_{C_9})$	-	-2	0.5

The process has 42 measurement candidates (Table 2), which include (1) concentrations of chemical species inside the reactors, liquid-liquid separator, and in the feed streams, (2) temperatures in the reactors, and (3) volumetric flowrates in the feed streams.

The manipulated variables are the volume flowrates of the feed streams containing reactants: q_i, $i = \{1, 2, 3, 4\}$. Feed streams in the third mixer containing solvents are assumed to remain constant with nominal values of $q_5 = 0.2$ mL/min and $q_6 = 0.5$ mL/min. The volumetric flowrates need to satisfy the constraints:

$$0 \leq q_i \leq 4 \text{ mL/min} \quad \text{for } i = \{1, 2, 3, 4\}. \tag{18}$$

Parametric model uncertainty is considered in the separation coefficient of atropine D_{C_9} in the liquid-liquid separator and in the reaction kinetic parameters (pre-exponential factors k_i and activation energies E_{Ai}). Disturbances are the molarity of components C_1 and C_7 in the feed streams and the reactor temperatures. The parametric uncertainties are assumed to be normally distributed with mean and standard deviations (Table 3). Sensor noise is modeled as Gaussian noise with zero mean and standard deviations of 2.5% of the nominal value for volume flowrates and concentrations, and 1 K for temperatures.

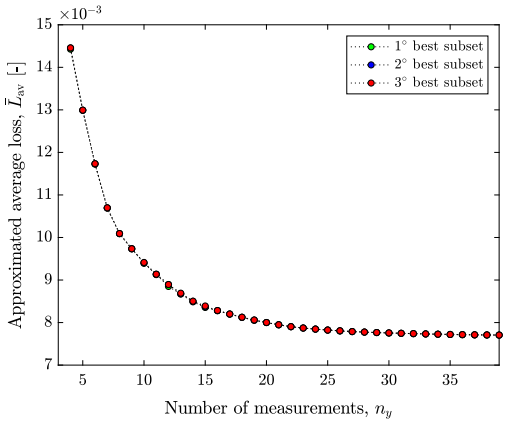

Fig. 3. Approximated average loss as a function of the number of measurements.

The objective function is to maximize is the E-factor, which is the mass of waste per mass of product.

4. RESULTS AND DISCUSSION

4.1 Screening of self-optimizing controlled variables

The best 3 measurement subsets with the least losses in increasing order were computed for each number of measurements, ranging from 4 to 42, by using the mixed integer quadratic programming formulation in Section 2.2. The optimization was solved using the `cplexmiqp` function in `CPLEX v12.8.0`.

The Pareto frontier corresponding to the approximated average loss as a function of the number of measurements in shown in Fig. 3. The indices of the optimal measurement subsets are listed in Table 4.

4.2 Steady-state validation using the nonlinear model

The measurement subsets screened by the MIQP algorithm were evaluated using the original steady-state nonlinear model. Each set of controlled variables was tested in closed-loop simulations over 100 scenarios of normally distributed disturbances and sensor noise realizations, with mean and standard deviations in Section 3. The average losses as a function of the number of measurements included in the control structure are in Fig. 4. The measurement subsets with the least average losses are in Table 5.

Deciding the number of measurements to include in the control structure is a tradeoff between the reduction in the average loss and the increase in structure complexity and investment cost that comes with the addition of more sensors. From the average losses shown in Fig. 4, a control

Table 4. Measurement subsets with the least approximated average losses.

n_y	Measurement subset	$\bar{L}_{\text{av}}$ ($\times 10^{-3}$)
4	[13, 19, 20, 26]	14.427
	[3, 13, 20, 26]	14.456
	[11, 13, 20, 26]	14.457
5	[13, 17, 19, 20, 36]	12.990
	[11, 13, 17, 20, 36]	12.992
	[3, 13, 17, 20, 36]	12.994
6	[7, 13, 19, 20, 21, 26]	11.725
	[7, 11, 13, 20, 21, 26]	11.733
	[3, 7, 13, 20, 21, 26]	11.734
7	[7, 13, 19, 20, 21, 24, 26]	10.690
	[7, 11, 13, 20, 21, 24, 26]	10.696
	[3, 7, 13, 20, 21, 24, 26]	10.698
⋮	⋮	⋮
42	all measurements	7.7056

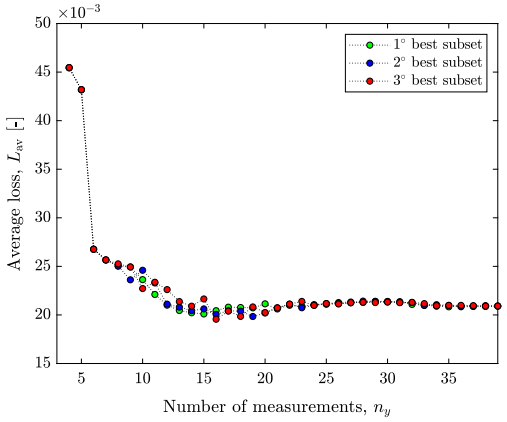

Fig. 4. Average losses as a function of the number of measurements for 100 scenarios of normally distributed disturbances and sensor noise realizations.

Table 5. Subsets of measurements selected by the global-self-optimizing control algorithm.

n_y	Measurement subset	L_{av} ($\times 10^{-3}$)
4	[13, 19, 20, 26]	45.428
5	[13, 17, 19, 20, 36]	43.177
6	[7, 13, 19, 20, 21, 26]	26.753
7	[7, 13, 19, 20, 21, 24, 26]	25.654
8	[7, 13, 17, 19, 20, 21, 24, 36]	25.033
9	[7, 12, 13, 17, 20, 21, 23, 24, 36]	23.614
10	[7, 12, 13, 17, 20, 21, 22, 23, 24, 36]	22.717
⋮	⋮	⋮
42	all measurements	20.889

structure composed of 6 measurements gives a reasonable tradeoff between these two objectives. The combination matrix associated with this control structure is:

$$\mathbf{H} = \begin{bmatrix} -1.65 & -1.39 & 0.29 & -69.9 & 2.88 & 1.10 & 2.15 \\ -2.15 & -1.57 & 0.79 & -38.9 & 15.5 & 1.32 & 1.87 \\ -1.90 & -1.21 & 1.55 & -8.44 & 10.6 & 0.98 & 0.80 \\ -3.18 & -2.25 & 1.65 & -12.9 & 15.0 & 1.84 & 2.45 \end{bmatrix} \quad (19)$$

Figure 5 shows that the self-optimizing control structure composed of 6 measurements (in blue) achieves near-zero losses compared with an open-loop policy of keeping

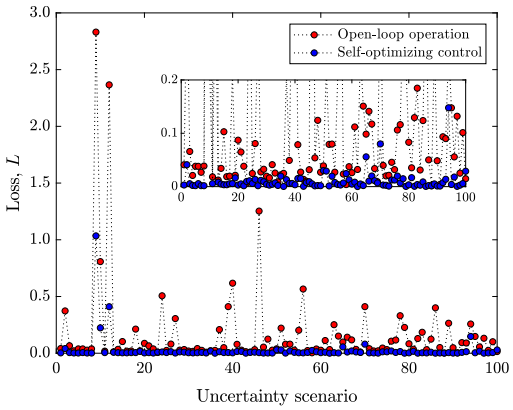

Fig. 5. Losses for 100 scenarios of normally distributed disturbances and sensor noise realizations obtained by self-optimizing control (blue) and open-loop operation (red).

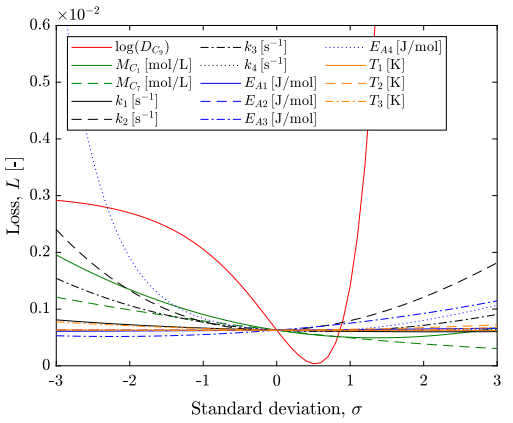

Fig. 6. Sensitivity of the loss for different disturbances.

the nominally optimal value of the manipulated variables constant in the presence of disturbances and sensor noise.

The disturbance with the highest impact on the loss is the parametric model uncertainty in the distribution coefficient of atropine, followed by the activation energy of the fourth reaction which controls the production of apoatropine (undesired byproduct). This relative importance is illustrated in Fig. 6, which shows the loss as a function of the standard deviation of the different disturbances.

4.3 Dynamic simulation

The proposed control structure was validated using closed-loop dynamic simulations. Initially, the system is operated at the nominal conditions. Subsequently, for every 1000 min, the operating conditions in the plant change due to the impact of disturbances. The three disturbance scenarios were generated by Monte Carlo sampling given the standard deviations in Table 3. In the simulation, normally distributed sensor noise is assumed with zero mean and standard deviations of 2.5% of the nominal value for the concentration measurements.

Decentralized PI controllers were used to track the self-optimizing controlled variables to a constant setpoint of

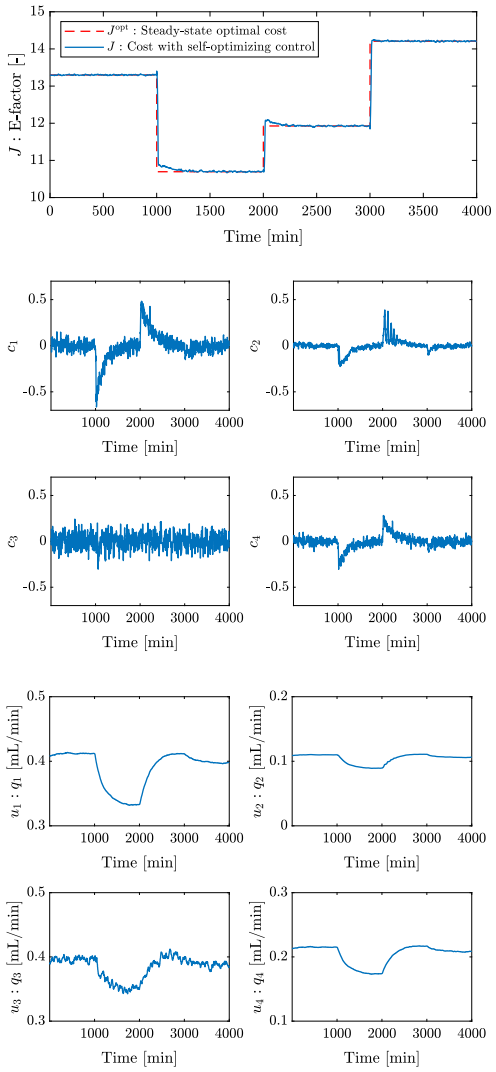

Fig. 7. Dynamic closed-loop simulation of the self-optimizing control structure in the presence of disturbances and sensor noise. J is the cost function (E-factor), c_i are the controlled variables, and u_i are the manipulated variables.

zero. The pairing was based on the steady-state relative gain array (RGA) and tuning based on the SIMC method.

Figure 4.3 shows the cost function, controlled variables, and manipulated variables in the closed-loop dynamic simulation. The cost obtained by tracking the self-optimizing controlled variables to constant setpoints followed very closely the optimal steady-state cost for each disturbance scenario without the need to reoptimize the setpoints when disturbances occur.

5. CONCLUSION

This article demonstrates the use of self-optimizing control in a continuous pharmaceutical manufacturing plant. Uncertainty is handled by controlling a carefully chosen combination of measurements to a constant setpoint. The method was applied to a simulated plant based on the continuous-flow synthesis of atropine. A mixed integer quadratic program based on a global self-optimizing control algorithm was solved to find the best self-optimizing variables. The closed-loop simulations indicate that this control policy is able to mantain the process close to optimality despite disturbances, sensor noise, and parametric model uncertainty.

ACKNOWLEDGEMENTS

DPP acknowledges financial support from "La Caixa" Foundation during his research stay at MIT. The work of AN and RDB was supported by the DARPA Make-It program under contract ARO W911NF-16-2-0023.

REFERENCES

Adamo, A., Beingessner, R.L., and et al. (2016). On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science*, 352(6281), 61–67.

Jäschke, J., Cao, Y., and Kariwala, V. (2017). Self-optimizing control–A survey. *Annu. Rev. Control*, 43, 199–223.

Kariwala, V. and Cao, Y. (2010). Bidirectional branch and bound for controlled variable selection. Part III: Local average loss minimization. *IEEE T. Ind. Inform.*, 6(1), 54–61.

Kariwala, V., Cao, Y., and Janardhanan, S. (2008). Local self-optimizing control with average loss minimization. *Ind. Eng. Chem. Res.*, 47(4), 1150–1158.

Morari, M., Arkun, Y., and Stephanopoulos, G. (1980). Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. Process decomposition and the classification of the control tasks. analysis of the optimizing control structures. *AIChE Journal*, 26(2), 220–232.

Myerson, A.S., Krumme, M., Nasr, M., Thomas, H., and Braatz, R.D. (2015). Control systems engineering in continuous pharmaceutical manufacturing. *J. Pharm. Sci.*, 104(3), 832–839.

Nikolakopoulou, A., von Andrian, M., and Braatz, R.D. (2019). Plantwide control of a compact modular reconfigurable system for continuous-flow pharmaceutical manufacturing. In *American Control Conf.*, 2158–2163.

Nikolakopoulou, A., von Andrian, M., and Braatz, R.D. (2020). Fast model predictive control of startup of a compact modular reconfigurable system for continuous-flow pharmaceutical manufacturing. In *American Control Conf.* In press.

Powell, W.B. (2019). A unified framework for stochastic optimization. *European Journal of Operational Research*, 275(3), 795–821.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control*, 10(5), 487–507.

Ye, L., Cao, Y., and Yuan, X. (2015). Global approximation of self-optimizing controlled variables with average loss minimization. *Ind. Eng. Chem. Res.*, 54(48), 12040–12053.

Yelchuru, R. and Skogestad, S. (2012). Convex formulations for optimal selection of controlled variables and measurements using mixed integer quadratic programming. *Journal of Process Control*, 22(6), 995–1007.