CONTROL STRUCTURE DESIGN TO ACHIEVE MULTIPLE PERFORMANCE CRITERIA

Yongsong Cai and Thomas E. Marlin*

Department of Chemical Engineering, McMaster University
1280 Main Street West, Hamilton, Ontario, Canada, L8S 4L7

Abstract: The control structure has the dominant effect on control performance. Developing systematic methods for selecting the structure is challenging because of the multi-criteria nature of the problem. Some criteria include dynamic behavior of the controlled and manipulated variables, equipment capacities, and loop integrity. In this paper, a method is presented to address the complex nature of the problem, while using a formulation that can be solved in a reasonable computing time. The primary criterion is controlled variable performance, but additional criteria such as integrity are considered using meaningful metrics, such as Relative Gain Array. The crucial importance of using realistic scenarios and not relying on heuristics is demonstrated using two process examples.

Keywords: control structure selection, decentralized control, plantwide control, complementarity formulation

1. INTRODUCTION

Control structure selection is an important decision at the beginning of the control system design. Available control design methods range from heuristic rules to approximate metrics to mathematical programming. Luyben, *et al.* (1998) presented a nine-step heuristic design procedure for plantwide control design. Bristol (1966) developed the most widely used metric, the Relative Gain Array (RGA), which only needs steady state gain and gives very useful information for integrity and robustness. A more complete analysis of control performance requires the transient response of key variables. Narraway and Perkins (1994) formulated the control structure selection and controller tuning together as a dynamic optimization using MINLP. Wang and McAvoy (2001) chose the control structure by optimizing dynamic performance for several subsets of controlled variables sequentially. Kookos and Perkins (2002) used linear dynamic model and integer variable for both structure and tuning selection, which makes a MILP problem with a very large number of integer variables.

This paper presents a control structure design method that extends the optimization of the time-domain transient response and gives the best overall behavior in industrial situations, even though the design might violate one or more common heuristics. Since control design is naturally a multiple criteria decision-making procedure, our method is flexible enough to consider several criteria simultaneously. In addition, the design must be performed for a set of conditions, i.e., a scenario, which represents the realistic situation. For example, the scenario should include realistic disturbances, noise, uncertainty and equipment bounds. Finally, since good controller tuning is essential when comparing alternative designs, our method includes a tuning optimization. To achieve these formidable results within reasonable computing time, we restrict ourselves to linear dynamic models and employ an efficient new optimization algorithm.

The paper is presented as follows; Section 2 describes the approach used for multiple criteria optimization, Section 3 defines the formulation for realistic scenarios, Section 4 discusses special

*Author to whom all correspondence should be addressed
Email: marlint@mcmaster.ca

aspects of the solution method, and Section 5 presents results from two cases studies.

2. MULTIPLE CRITERIA

There are many control design objectives (Marlin, 2000) that address a range of requirements including safety, smooth operation, product quality, and production rate. In addition, most systems involve inequality constraints on variables; some of these are hard (e.g., equipment performance limits) and others are soft (e.g., deviation from set point). The most often used method to formulate multiple criteria optimization is to penalize the weighted summation of all criteria. In spite of its mathematical simplicity, this method has a major drawback. The weighting should reflect the relative importance among criteria, but in most cases we do not know this sensitivity information before we solve the problem.

For multi-criteria optimization, we choose the e-constraint method (Steuer, 1986). It optimizes one criterion with the other criteria bounded as constraints.

$$\begin{aligned} &\min_{\mathbf{x}} c_j(\mathbf{x}) \\ &s.t.\quad c_i(\mathbf{x}) \le e_i \quad \forall i \ne j \\ &\qquad \mathbf{h}(\mathbf{x}) = 0 \\ &\qquad \mathbf{g}(\mathbf{x}) \le 0 \end{aligned} \tag{1}$$

where $c_j(\mathbf{x})$ is the criterion as objective function and $c_i(\mathbf{x})$ are the criteria bounded by e_i. This method can be solved by standard optimization software and the result is easy to understand. Steuer (1986) points out that by solving the problem several times with different e_i, we can get a very useful byproduct, sensitivity among different criteria.

$$\frac{\partial c_j(\mathbf{x})}{\partial e_i} \tag{2}$$

Our major focus in this paper is dynamic performance; therefore, a measure of performance of the controlled variable deviation from set point is very natural choice for the objective function, c_j. Many other criteria could be included as the ancillary objectives, c_i. In this paper, we will demonstrate the approach by including integrity as an optional bound. Integrity has many definitions; here, we select "a system has integrity if the closed-loop dynamic system is stable when one or more loops are placed in manual (off) without changing the sign of the active feedback controller gains". One widely used metric is relative gain array (RGA), which is a necessary condition for good integrity.

3. FORMULATION WITH REALISTIC SCENARIO

Control performance is strongly affected by the scenario. Certainly, large, fast disturbances are worse than small, slow disturbances. However, many other characteristics are important in control design and many of these have been ignored in previously published methods. To achieve the flexibility required, we propose an optimization formulation that involves a linear dynamic model including a saturation model of the manipulated variable. The decision variables are the integer variables that specify the control structure and the controller tuning. The objective function is the integral error of the controlled variable deviation with a penalty for the control actions. The resulting formulation is given in the following equations.

$$\min_{\delta_{ij}, K_c, K_I, s_k^U(t), s_k^L(t)} \sum_k \sum_{t=1}^{tf} [\mathbf{Q} \cdot \mathbf{ae}_k(t) + \mathbf{R} \cdot \mathbf{au}_k(t)] \tag{3}$$

$s.t.$

$$\begin{aligned} \mathbf{x}_k(t+1) &= \mathbf{A}_k \mathbf{x}_k(t) + \mathbf{B}_k \mathbf{u}_k(t) + \mathbf{W}_k \mathbf{d}(t) \\ \mathbf{y}_k(t) &= \mathbf{C}_k \mathbf{x}_k(t) + \mathbf{V}_k \mathbf{d}(t) + N(t) \end{aligned} \tag{4}$$

$$\begin{aligned} \Delta\mathbf{u}_k(t) &= \mathbf{K}_c\left[\mathbf{e}_k(t) - \mathbf{e}_k(t-1) + \mathbf{K}_I \mathbf{e}_k(t)\right] \\ \mathbf{e}_k(t) &= \mathbf{sp}(t) - \mathbf{y}_k(t) \end{aligned} \tag{5}$$

$$\begin{aligned} &\mathbf{u}_k(t) = \mathbf{u}_k(t-1) + \Delta\mathbf{u}_k(t) - \mathbf{s}_k^U(t) + \mathbf{s}_k^L(t) \\ &s_{k,i}^U(t) \ge 0 \quad complements \quad u_{k,i}(t) \le u_{U,i} \\ &s_{k,i}^L(t) \ge 0 \quad complements \quad u_{k,i}(t) \ge u_{L,i} \end{aligned} \tag{6}$$

$$\begin{aligned} &-\mathbf{ae}_k(t) \le \mathbf{e}_k(t) \le \mathbf{ae}_k(t) \\ &-\mathbf{au}_k(t) \le \Delta\mathbf{u}_k(t) \le \mathbf{au}_k(t) \\ &\mathbf{ae}_k(t) \ge 0, \quad \mathbf{ae}_k(t) \ge 0 \end{aligned} \tag{7}$$

$$\begin{aligned} &\text{-}M_{ij}\delta_{ij} \le K_{c,ij} \le M_{ij}\delta_{ij} \\ &\text{-}M_{ij}\delta_{ij} \le K_{I,ij} \le M_{ij}\delta_{ij} \end{aligned} \tag{8}$$

$$\sum_i \delta_{ij} = 1, \quad \sum_j \delta_{ij} = 1 \tag{9}$$

$$\sum_j \lambda_{ji}\delta_{ij} \ge 0 \qquad \forall i \tag{10}$$

We use the summation of absolute error $\mathbf{ae_k}$(t) and absolute manipulated variable movement $\mathbf{au_k}$(t) as objective function in (3). Equation (7) calculates the absolute variable of error and manipulated variable movement that are used in objective function. The plant is described by a linear time-invariant state space model in. (4), and we use more than one model to represent model mismatch. In (4), subscription k describes different models, N(t) is measurement noise, $\mathbf{d}$(t) is disturbance. Multi-loop PI controllers are used, (5). The tuning parameters are controller gain, $\mathbf{K}_c$; integral gain, $\mathbf{K}_I$, which equals $\Delta T/\mathbf{T}_I$. Equation (6) expresses the capacity of the manipulated variable with upper and lower bounds, $u_{U,i}$ and $u_{L,i}$.

The slack variables, $\boldsymbol{s}^U_k$(t) and $\boldsymbol{s}^L_k$(t), express the amount that the controller output violates the bound on the manipulated variable. For instant, if the calculated manipulated variable output $u_{k,i}$(t-1)+$\Delta u_{k,i}$(t) is greater than its upper bound $u_{U,i}$, the actual manipulated variable output $u_{k,i}$(t) is clamped to its upper bound by allowing slack variable $s^U_{k,i}$(t) to be nonzero. The complementarity constraints guarantee that slack variables become nonzero only when the manipulated variables reach their limits. For instant, if manipulated variable $u_{k,i}$(t) is less than its upper bound $u_{U,i}$, that means inequality $u_{k,i}(t) \le u_{U,i}$ is not active, and by the complementarity constraints, the slack variable $s^U_{k,i}(t)$ must be zero. We note that this formulation also prevents reset-windup.

δ is a binary matrix that has the same size as the controller and defines loop pairing. We define $\delta_{ij} = 1$ if the controller pairs u_i with y_j and $\delta_{ij} = 0$ otherwise. Equation (8) states that the controller tuning parameters, $K_{c,ij}$ and $K_{I,ij}$ can be nonzero only when the corresponding δ_{ij} is one, which means the pairing $u_i - y_j$ is chosen. Equation (9) limits the possible structure to diagonal structure, which means that there is one and only one controller in each row and column.

Equation (10) forces the pairing we choose to have a positive RGA value, which is a necessary condition for integrity. As we will see in the case studies, enforcing integrity has a major effect on the control performance. Therefore, we will solve the problem with and without (10) to determine whether the good integrity property costs excessive degradation in control performance.

4. SOLUTION METHOD

Equations (3)-(10) define a mixed integer, non-convex, non-linear programming problem, which is challenging to solve. We take several steps to simplify the solution. First, we evaluate the relative gain at each node of the integer problem, i.e., at each pairing before optimizing the continuous problem. This procedure is appropriate because the (steady-state) relative gain depends on the process and the pairing, but not the disturbance or controller. To address the non-convexity of the continuous problem, we take two actions.

4.1 Good Initial Points

The optimal tuning problem is non-convex, so a good starting point is very important to solve the problem. Since we can choose a pairing with a negative RGA, some of the controller signs might need to be different from their SISO loop tuning. If a specific pairing has a negative relative gain, this does not ensure that the controller for this pairing should have its gain inverted. Clearly, the non-convex problem is complicated and covers a very wide range of tuning values. We propose a two-step procedure: a region elimination followed by an NLP in the identified region. The regions are defined by the controller gain (sign and magnitude) and the integral constant (magnitude). We have observed that good multiloop control can be achieved with an aggressive controller with a large $K_{c,ij}$ or a large $K_{I,ij}$ (small $T_{I,ij}$), depending on the particular problem. The regions are summarized in table 1.

We evaluate the control performance within each of these regions for the design scenario by simulation. While the table can contain many cases, each case is solved quickly. Then, we bound the tuning within the best region for the control design optimization problem.

Table 1 Possible tuning combinations

K_{c1}	K_{I1}	K_{c2}	K_{I2}	…
+low	low	+low	low	…
+high	low	-low	low	…
-low	high	+low	low	…
…	…	…	…	

4.2 IPM Solver – IPOPT-C

A major challenge remains because the valve saturation formulation (6) introduces the complementary constraints. In our experience, standard NLP solvers will not reliably solve this problem. Therefore, we have selected to use IPOPT-C, which is an Interior Point Method (IPM) solver solves general NLP. It has been extended to handle complementary constraints (Raghunathan and Biegler, 2003). The central-path following feature of this solver offers the potential for improved performance for this non-convex optimization. Good results have been obtained, but global optimization cannot be guaranteed.

5. CASE STUDIES

5.1 Fluidized Catalytic Cracker (FCC)

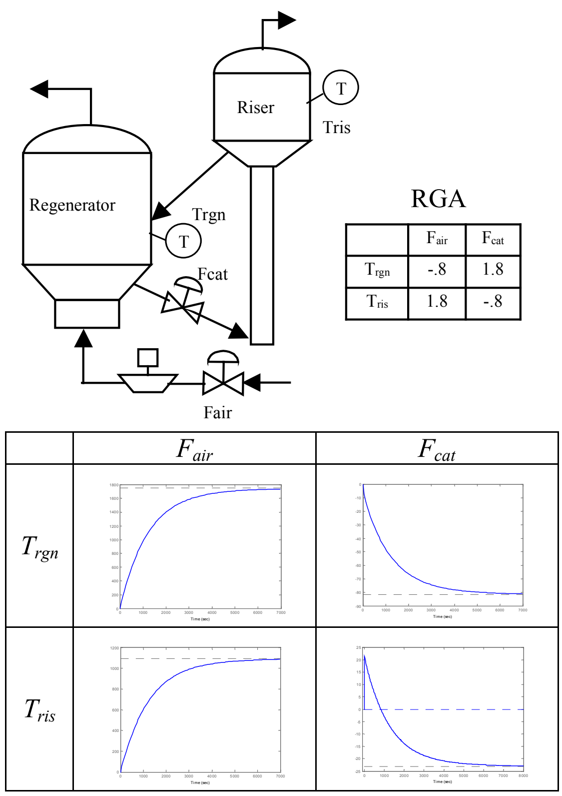

Fig. 1. Fluidized catalytic cracker process and open-loop response

We will apply the method to the loop pairing for the fluidized catalytic cracker unit shown in Fig. 1. It converts heavy oil into lighter and more valuable products, such as gasoline and fuel oil. The reaction section has two major units, one is a plug flow (transportation) reactor called the riser, which has a residence time of only a few seconds. The other unit is a fluidized bed vessel called regenerator, in which the catalyst is regenerated by burning the coke produced in the reactor using air. The inventory of catalyst in a typical regenerator is on the order of 60 tons. Two temperatures must be controlled to achieve smooth operation. The temperature of the riser (T_{ris}) is directly related to product quality and yield and to temperature limits of key equipment; therefore, it should be tightly controlled. The temperature of the regenerator (T_{rgn}) is not as crucial as riser, but it should be controlled in a range about its set point. Two available manipulated variables are flow of catalyst (F_{cat}) and flow of air (F_{air}).

The linearized model used in this study is based on Arbel *et al.* (1996). The open-loop response (Fig. 1) shows that the system has slow dynamics except loop T_{ris}–F_{cat}, which has a fast inverse response. The RGA for this 2x2 system is shown in Fig. 1. If we want the control system with integrity, we have only one choice, pairing on T_{rgn}–F_{cat}, T_{ris}–F_{air}. The other pairing will have poor integrity. Many heuristic design approaches would dictate that the relative gain must be positive. However, we will proceed with a systematic evaluation of the dynamic performance. The evaluation will begin with the simplest (unrealistic) scenario and add additional factors in the scenario until a realistic definition is attained (Table 2). Using this procedure, we will be able to follow the effects of the scenario on the control design.

Table 2 Cases for different scenarios

Case	A	B	C	D	E
RGA	+	+	+	+	-
Mismatch		×	×	×	×
Noise			×	×	×
Capacity				×	×
$K_p * K_c \approx$	10^6	100	1	1	1
IAE(T_{ris})	21.2	54.3	32.7	72.2	31.1

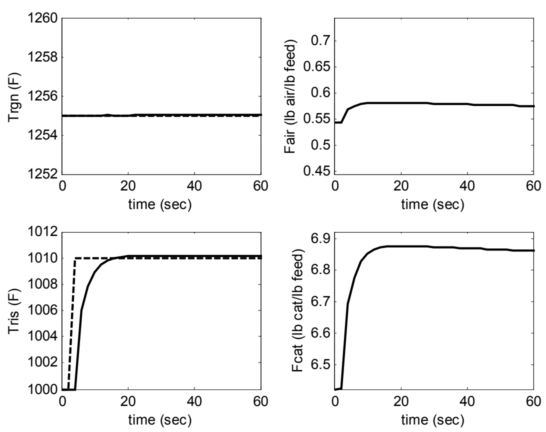

Fig. 2. Case A

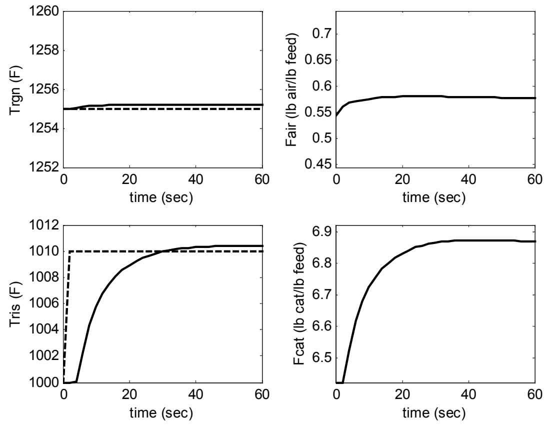

Fig. 3. Case B

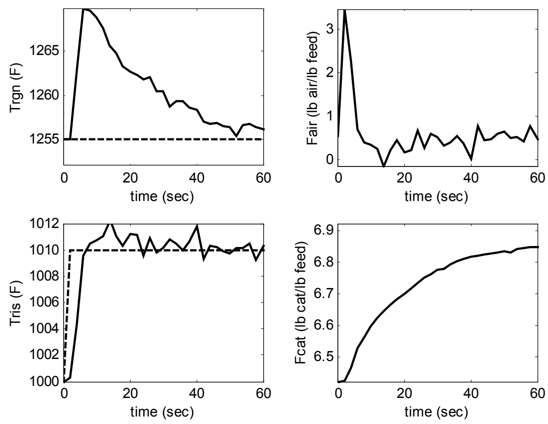

Fig. 4. Case C

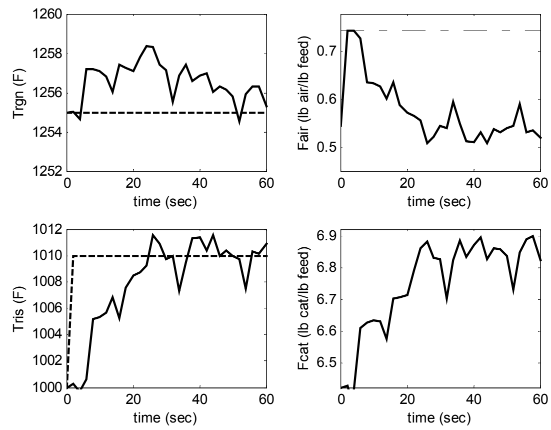

Fig. 5. Case D

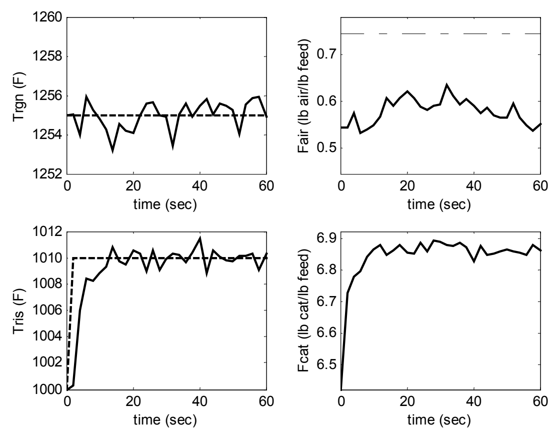

Fig. 6. Case E

We begin with the simplest scenario, without mismatch model, noise and valve constraints. The performance looks very good (Fig. 2). One guideline for tuning is that the product of the process gain K_p times the controller gain K_c should be around one. But, one of the controllers has a value of about 10^6! If we don't consider model mismatch, this tuning gives good performance, but a very small model mismatch will make the closed-loop system unstable. Fig. 3 shows the result for optimizing the base case and mismatch models simultaneously, with one tuning. The controller gain is significantly reduced but still very big. T_{rgn} only moves about 0.5°F, which will be overwhelmed by noise in real plant. The big controller gain will amplify the noise and make the performance much worse than its prediction. Fig. 4 shows the result with noise added in the formulation. The tuning now agrees with our guideline, and the performance is good. But in order to achieve this good performance, F_{air} has to make a big initial increase, which would require an air blower with about 600% extra capacity. This extra capacity would involve a two million dollar increase in plant capital investment. Fig. 5 shows the performance with a realistic extra plant capacity, which is the formulation we propose, which involves equations (3) – (10). The tuning and transient becomes reasonable, and the dynamic performance is realistic. In Fig. 5 the trajectory of F_{air} stays at its upper bound for several steps, which is saturation handled properly with the formulation. We note that the performance of the riser temperature has degraded in Fig. 5 and is likely not adequate for this critical variable.

Therefore, we chose to relax the requirement that the relative gain be positive. We report the results for the most demanding scenario in Fig 6, which shows the excellent performance for the riser temperature. The IAE with the negative RGA pairing is only half of the value achieved with the positive RGA pairing.

The final decision on the control structure is made by the engineer based on the rigorous optimization results. For FCC process, the temperature of riser is crucial, since it is directly related to profit and equipment protection. We want to keep it as close to its set point as possible. If the regenerator controller is placed in manual, the control system would be unstable because lack of integrity. However, the regenerator temperature would drift away slowly because of its huge volume, so the operators have time to respond. Based on this analysis we would choose dynamic performance and give up integrity; we choose negative RGA pairing, which is T_{rgn}–F_{air}, T_{ris}–F_{cat}. This is the industry standard control design for the FCC process considered in this example (Arbel, *et al.*, 1996). Note that it violates a commonly cited control heuristic, but it is widely applied.

5.2 Fired Heater

This case involves the control of the fired heater in Fig. 7, which has four burners and four coils in one firebox. The control goal is to manipulate the fuel to the four burners to control the four coil outlet temperatures. The transfer function model can be found in Manousiouthakis *et al.* (1986). The dynamics of each input-output is first-order, and the system has strong interaction. Equation (11) reports the relative gain array.

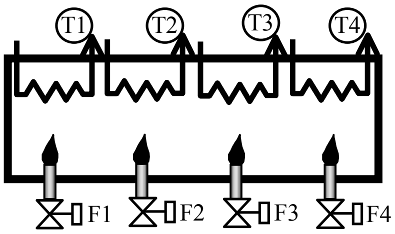

Fig. 7. Fired heater

$$RGA = \begin{bmatrix} \mathbf{1.748} & -0.686 & -0.096 & 0.034 \\ -0.727 & \mathbf{1.874} & -0.092 & -0.055 \\ -0.055 & -0.092 & \mathbf{1.874} & -0.727 \\ 0.034 & -0.096 & -0.686 & \mathbf{1.748} \end{bmatrix} \quad (11)$$

We see that only two pairings have positive RGA and good integrity. However, the previous case demonstrated the danger of basing a design on a heuristic. Therefore, we proceed to apply additional short-cut estimates of control performance that can be obtained from steady-state gain information. Here, we use the relative disturbance gain, RDG (Stanley et al, 1985). When the value of the RDG is less than 1.0, the multiloop performance, as measured by the integral error, is better than the single-loop performance.

$$RDG = \begin{bmatrix} 0.5479 & 0.2397 & 0.1027 & \mathbf{0.1096} \\ 0.3288 & 0.3425 & \mathbf{0.1370} & 0.1918 \\ 0.1918 & \mathbf{0.1370} & 0.3425 & 0.3288 \\ \mathbf{0.1096} & 0.1027 & 0.2397 & 0.5479 \end{bmatrix} \quad (12)$$

We see that all pairings are predicted to be better than single-loop, which indicates that the individual loops interact in a favorable manner. The highlighted pairings have the best (lowest) integral error. However, it is well known that the integral error can be small because of positive and negative cancellation (this is not the squared error).

Therefore, we proceed to the proposed design technique. We can evaluate the best performance with integrity (positive RGA pairings) and without the requirement for positive RGA pairings. In this case, the method finds the best dynamic control performance for a realistic scenario is achieved by the diagonal design highlighted in (11), see Fig. 8.

These rigorous results do not conform to the prediction from the RDG; in fact, the design indicated by (12) gives one of the poorest dynamic performances (Fig. 9). The reason is the complex feedback dynamics that occur with this pairing, which results in poor performance. This poor performance is not reflected in the heuristic, which is not affected by large positive and negative oscillations in the controlled variable that cancel. Again, we see that heuristics and approximate metrics can be misleading. We note that the RDG has proven useful in eliminating candidates in other studies, so that this case study does not generally invalidate the RDG metric, which should always be combined with a transient analysis before a design is finally selected.

The choice for the engineer in this case is easy, because good dynamic and integrity are achieved simultaneously. The diagonal controller highlighted in (11) is selected.

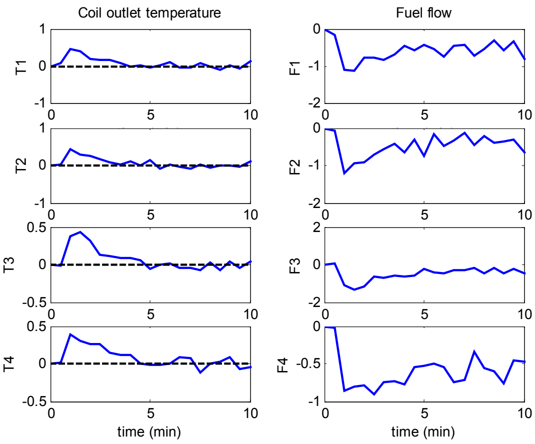

Fig. 8. Disturbance response of the pairing highlighted in (11).

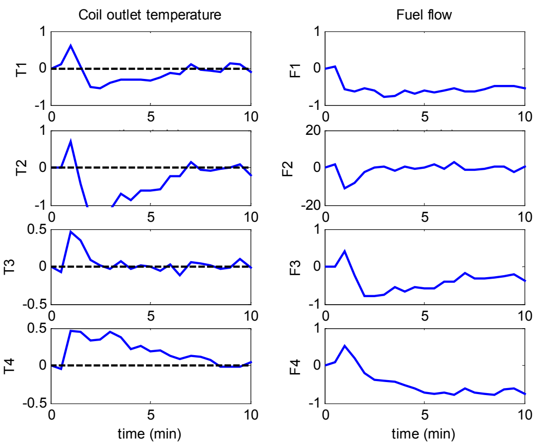

Fig. 9. Disturbance response of the pairing highlighted in (12).

6. CONCLUSION

This paper has discussed an optimization-based control structure design method. We demonstrate that a realistic scenario formulation is essential; this method includes important factors such as model mismatch, measurement noise and equipment capacity limits, to makes the resulting tuning and dynamic performance prediction close to the achievable performance in real. The multiple criteria framework enables us to tradeoff different design objectives, such as dynamic performance and integrity. We anticipate its application as an interactive procedure. Engineers can use their process knowledge to make the final decision based on results with different limits on the inequality constraints in (1).

Future work involves extending this methodology to handle more complex control structures, such as such a block-centralized structure.

REFERENCES

Arbel, A., R.H. Rinard and R. Shinnar (1996). Dynamics and control of fluidized catalytic crackers 3. Designing the control system: Choice of manipulated and measured variables for partial control. *Ind. Eng. Chem. Res.,* **35**, 2215-2233.

Bristol, E.H. (1966). On a new measure of interaction for multivariable process control. *IEEE Trans. Autom. Control*, **11**, 133-134.

Kookos, I.K. and J.D. Perkins (2002). Regulatory control structure selection of linear systems. *Compu. & Chem. Eng.*, **26**, 875-887.

Luyben, M.L., B.D. Tyreus and M.L. Luyben (1998). *Plantwide Process Control*. McGraw-Hill, New York.

Manousiouthakis,V., R.Savage and Y. Arkun (1986). Synthesis of decentralized process control structures using the concept of block relative gain. *AIChE Journal,* **32**, 991-1003.

Marlin, T.E. (2000). *Process control – designing process and control system for dynamic performance*, 2nd Ed, Chap 24. McGraw-Hill, New York.

Narraway, L. and J.D. Perkins (1994). Selection of process control structure based on economics. *Compu. & Chem. Eng.*, **18**, S511-S515

Raghunathan, A.U. and L.T. Biegler (2003). Mathematical programs with equilibrium constraints (MPECs) in process engineering. *Compu. & Chem. Eng.*, **27**, 1381-1392.

Stanley, G., M. Marino-Galarraga and T.J. McAvoy (1985). Short-cut operability analysis: 1. The relative disturbance gain. *Ind. & Eng. Chem. Proc. Des. Dev.* **24**, 1181-1188.

Steuer, R.E. (1986). *Multiple Criteria Optimization: Theory, Computation, and Application*. John Wiley & Sons, New York.

Wang, P. and T. McAvoy (2001). Synthesis of plantwide control systems using a dynamic model and optimization. *Ind. & Eng. Chem. Res.* **40**, 5732-5742.