Optimal Control of Catalytic Methanol Conversion to Formaldehyde

A. Faliks,[†] R. A. Yetter,[‡] C. A. Floudas,[§] S. L. Bernasek,[†] M. Fransson,[†] and H. Rabitz*,[†]

Departments of Chemistry, Mechanical and Aerospace Engineering, and Chemical Engineering, Princeton University, Princeton, New Jersey 08544

Received: March 13, 2000; In Final Form: December 18, 2000

An optimal control methodology is applied to find the heat and oxygen flux profiles, distributed along the length of a plug flow reactor, for the conversion of methanol to formaldehyde. The calculations use models for the gas-phase and catalytic [$MoO_3-Fe_2(MoO_4)_3$] reactions. The reactor designs show that a distributed heat flux improves formaldehyde yields, but an oxygen flux does not affect the results. Formaldehyde mass fractions of over 90% have been achieved in the simulations. The solutions obtained, although not proven to be globally optimal, are of very high quality. A fully nonlinear robustness analysis of the formaldehyde production with respect to the catalyst model variables is performed by the use of a high dimensional model representation. This representation is similar to the ANOVA decomposition used in statistics but does not require an increase in the number of data points as the dimensionality of the variable space increases. The most important variables are the catalyst surface area and the rate of formaldehyde desorption. The yield improvement from employing optimized fluxes is found to be robust to the catalytic model parameter values.

1. Introduction

Formaldehyde is an important chemical used for a wide variety of purposes. It may be synthesized by partially oxidizing methanol with air over a metal or metal oxide catalyst in a temperature range of 400−650 °C. Current catalysts are typically based on copper, silver, molybdenum alloy, or mixed iron and bismuth molybdates.

In this paper, we utilize an optimal control strategy to suggest ways to improve the yields of formaldehyde from methanol partial oxidation through the design of distributed oxygen and energy fluxes along the length of a plug flow reactor (PFR). Optimization of the fluxes is carried out using a combined gas-phase and catalytic reaction model. The gas-phase mechanism developed by Held and Dryer[1] consists of 22 species and 88 reactions. The model of the MoO_3-Fe_2 $(MoO_4)_3$ catalyst is based on kinetic data from Pernicone,[2] Liberti et al.,[3] Pernicone et al.,[4] and Batist et al.[5] None of the prior efforts at partial oxidation of methanol has attempted to optimize the distribution of energy and/or mass along the reactor length. Consequently, the yields from conventional reactor designs represent a lower bound on what might be achieved. The application of the optimal control methodology is very similar to that of Faliks et al.[6−8] where high-quality solutions were obtained for the conversion of methane to acetylene and ethylene as well as the optimal control of free radical polymerization.

After the optimal control designs are achieved, a fully nonlinear robustness analysis is applied with respect to the parameters of the catalysis model. A random sampling-high dimensional model representation (RS-HDMR) technique is used for this purpose; it is similar to the analysis of variance (ANOVA) decomposition used in statistics. The nonlinear HDMR analysis is needed because the high uncertainty characteristic of the catalysis parameters precludes the use of traditional linear sensitivity analysis based on small pertubations. A significant drawback of the ANOVA decomposition is the need to compute high-dimensional integrals over the parameter space requiring large numbers of model calculations, with the effort significantly growing as the dimensionality of the parameter space increases. The number of model runs needed for a RS-HDMR analysis is invariant with the dimensionality thereby producing a very efficient procedure.

In section 2, the reactor model is presented, and in section 3 the gas-phase and catalytic models are described with a base case set of kinetic parameters. Illustrative optimal control simulations with the base kinetic model are discussed in section 4. In this section it is shown that heat flux optimization results in an increased formaldehyde yield relative to a reference isothermal run. Section 5 presents the results of the nonlinear catalytic model robustness analyses. First, a robustness analysis is performed on the optimal formaldehyde yield with respect to broad scale conservative uncertainties in the parameter values of the base kinetic model. Then, the benefits of optimization in the distributed heat flux reactor are shown to be robust to the kinetic model parameters by separately optimizing 100 random systems, each with different values for the kinetic parameters. Finally, conclusions that consider the practical feasibility of the process are presented in section 6. The Appendices present the details of the optimal control algorithm and the RS-HDMR formulation.

2. Physical Formulation of the Flow Reactor

A plug flow reactor (PFR) is chosen as the basic reactor configuration. The PFR is a cylinder with constant cross sectional area and length L. Control is implemented through chemical and/or heat flux through the side wall of the reactor as a function of the position l along its length. The reactions are described by the production rate w_i of the ith species, $i = 1, ..., n$. The control variables are the fluxes of species i, denoted as j_i (mass/length−time), and the heat flux q (energy/length−

[†] Department of Chemistry.
[‡] Department of Mechanical and Aerospace Engineering.
[§] Department of Chemical Engineering.

10.1021/jp000951z CCC: $20.00
Published on Web 02/20/2001

time), as a function of position l. The mass fraction of species i in the reactor is denoted as $x_i(l)$, and the total mass flow rate is $F(l)$.

The following assumptions are made in modeling the PFR: (i) steady one-dimensional plug flow, (ii) instantaneous radial mixing, (iii) no diffusion along the axis of the reactor, and (iv) adiabatic reaction conditions. To make the assumptions realistic, the ratio of the length of the reactor to its radius is chosen to be greater than 25.

Since there are multiple sources of material flux entering the system (additionally, co-feed input is coincident with $j(0)$), the total mass balance for the flow rate is

$$F(l) = F(0) + \int_0^l \sum_{i=1}^{n} j_i \, \mathrm{d}l \tag{1}$$

By taking a differential control volume at position l and balancing the input and output mass and energy, we can arrive at the equations governing the composition in the reactor. Considering the conservation of mass, the species balance equation is

$$\begin{aligned} Fx_i + j_i \, \mathrm{d}l + w_i \, \mathrm{d}l &= \left(F + \frac{\mathrm{d}F}{\mathrm{d}l}\mathrm{d}l\right)\left(x_i + \frac{\mathrm{d}x_i}{\mathrm{d}l}\mathrm{d}l\right) \\ &= \left(F + \sum_{i=1}^{n} j_i \, \mathrm{d}l\right)\left(x_i + \frac{\mathrm{d}x_i}{\mathrm{d}l}\mathrm{d}l\right) \end{aligned} \tag{2}$$

The first term on the left represents the amount of the ith species flowing into the control volume. The second and third terms represent the amount of the species fluxed in from the side of the reactor and the amount produced or consumed in chemical reactions within the volume, respectively. Taking the infinitesimal limits of $\mathrm{d}x_i$ and $\mathrm{d}l$, we arrive at the mass conservation equation

$$\frac{\mathrm{d}x_i}{\mathrm{d}l} = \frac{1}{F}\left(w_i - x_i \sum_{k=1}^{n} j_k + j_i\right) \tag{3}$$

A similar approach is used to derive the energy conservation equation

$$\begin{aligned} T\sum_{i=1}^{n} C_{p_i} F x_i + \left(T_0 \sum_{i=1}^{n} C_{p_i} j_i + q\right)\mathrm{d}l - \sum_{i=1}^{n} H_{f_i} w_i \, \mathrm{d}l \\ = (T + \mathrm{d}T)\left[\sum_{i=1}^{n}\left(F + \frac{\mathrm{d}F}{\mathrm{d}l}\mathrm{d}l\right)\left(x_i + \frac{\mathrm{d}x_i}{\mathrm{d}l}\mathrm{d}l\right)C_{p_i}\right] \\ = \left(T + \frac{\mathrm{d}T}{\mathrm{d}l}\mathrm{d}l\right)\left[\sum_{i=1}^{n}\left(F + \sum_{i=1}^{n} j_i \, \mathrm{d}l\right)\left\{x_i + \left(\frac{1}{F}\left(w_i - x_i \sum_{k=1}^{n} j_k + j_i\right)\right)\mathrm{d}l\right\}C_{p_i}\right] \end{aligned} \tag{4}$$

where q is the heat influx, H_{f_i} is the heat of formation of species i, and C_{p_i} is the corresponding specific heat. The temperature of the influxed species is T_0. Taking the infinitesimal limits of $\mathrm{d}x_i$ and $\mathrm{d}l$ leads to the energy conservation equation

$$\frac{\mathrm{d}T}{\mathrm{d}l} = \frac{1}{F\sum C_{p_i} x_i}\left((T_0 - T)\sum_{i=1}^{n} C_{p_i} j_i - \sum_{i=1}^{n} H_{f_i} w_i - T\sum_{i=1}^{n} C_{p_i} w_i + q\right) \tag{5}$$

3. Chemical Reaction Model

The full model consists of combined gas-phase and catalytic reactions. A summary of the submodels is given below.

A detailed gas-phase mechanism was chosen for methanol oxidation[1] based on data from static reactor, flow reactor, shock tube and laminar flame experiments. The model covers conditions of temperature 300−2000 K, pressure 0.26−20 atm, and equivalence ratio 0.05−2.6. The basis of the gas-phase model is the subset of reactions involving hydrogen and oxygen and the intermediates and products associated with them. This submechanism determines the characteristics of the radical pool. The oxidation of carbon species begins with initiation reactions that are followed by radical attack on the methanol and production of small intermediates, and ends with a chain of steps:[9] aldehyde → CO → CO_2. The carbon monoxide/hydrogen/oxygen gas-phase submechanism used is taken primarily from Yetter et al.[10] The formaldehyde oxidation kinetics are based on the work of Hochgreb and Dryer[11] and Held and Dryer,[12] but some rates were changed to accurately reflect relative formaldehyde and methanol destruction rates in flow reactor experiments.[1] The methanol submechanism includes methoxy (CH_3O) and hydroxymethyl (CH_2OH). This submechanism involves reactions for initiation/decomposition, OH abstraction, reactions with H, and CH_3O/CH_2OH isomerization. Other reactions include formation of minor species such as formic acid (HCOOH) and 1,2-ethanediol (ethylene glycol) that act as intermediates. Also included is the creation of some C_2 hydrocarbon species due to the small amounts of methyl radicals that are created during the oxidation and pyrolysis of methanol and that may undergo recombination.

The catalytic model includes the kinetics of methanol oxidation over $MoO_3-Fe_2(MoO_4)_3$. The parameters used are based on flow reactor and pulse reactor studies in the temperature range 180−280 °C.[2−5] The proposed mechanism consists of several steps[2]:

(1) dissociative chemisorption of methanol;

(2) reduction of the catalyst, with formation of chemisorbed formaldehyde;

(3) water desorption, restoring sites for methanol chemisorption;

(4) formaldehyde desorption;

(5) catalyst reoxidation.

This process is represented by

$$CH_3OH + \square + O^{2-} \xrightarrow{k_1} (CH_3O - \square)^- + OH^- \tag{6}$$

$$(CH_3O - \square)^- + O^{2-} \xrightarrow{k_2} (CH_2O - \square)^{-2} + OH^- \tag{7}$$

$$2OH^- \overset{k_3}{\rightleftharpoons} H_2O + O^{2-} + \square \tag{8}$$

$$(CH_2O - \square)^{-2} \xrightarrow{k_4} CH_2O + \square^{2-} \tag{9}$$

$$\tfrac{1}{2}O_2 + \square^{2-} \xrightarrow{k_5} O^{2-} \tag{10}$$

where $\square$ is an anionic vacancy in the catalyst and O^{2-} is lattice oxygen. The reverse rate of reaction 8 is calculated through equilibrium arguments. This mechanism is similar to one proposed by Batist et al.[5] for the oxidative dehydrogenation of 1-butene over Bi molybdate catalysts. Water desorption has been shown to be reversible by infrared spectroscopy.[13] The rate parameters used are listed in Table 1, and these will be referred to as defining the base kinetic model. The results of Pernicone[2] indicate that the formaldehyde desorption is the rate-determining step while not excluding that the catalyst reduction is slow. These conclusions are supported by the work of Gorokhovatski

TABLE 1: Rate Parameters

$k_1{}^a = 1 \times 10^{11}$ L mol^{-1} s^{-1}
$k_2 = 1.9 \times 10^{10}$ L mol^{-1} s^{-1}
$k_3 = 1 \times 10^{11}$ L mol^{-1} s^{-1}
$k_4 = 10^{15} \exp[-16500/RT]^b$ s^{-1}
$k_5 = 1 \times 10^{11}$ atm$^{.5}$ s^{-1}

[a] The sticking coefficient is 0.5. [b] Activation energy units are calories.

and Evmenenko[14,15] and Popov et al.[16−19] In particular, Popov et al.[16−19] determined that neither methanol nor oxygen adsorption are the rate-determining steps. These authors also demonstrated that catalyst reduction is slower than reoxidation.

If the dissociative chemisorption of methanol is the rate-determining step, then the reaction rate should show a maximum versus the partial pressure of the reactants and the reaction rates should be equal in flow and pulse reactors. Pernicone et al.[4] reported that this is not the case. If, on the other hand, desorption is the slowest step, the reaction rate in the pulse reactor should be higher than the one in the flow reactor, as is indeed observed by Pernicone et al.[4] If formaldehyde desorption is the slowest step, the greatest part of the catalytic surface is occupied by formaldehyde in a flow reactor under stationary conditions. Therefore, even a large increase in the partial pressure of formaldehyde should cause only a slight decrease in the reaction rate. Pernicone et al.[4] have experimentally verified that this is true.

Liberti et al.[3] used a pulse microreactor to study the effects of temperature and flow rates on methanol oxidation over a $MoO_3-Fe_2(MoO_4)_3$ catalyst. Among their results is the activation energy of the formaldehyde desorption listed in Table 1. This activation energy has a small dependence on the flow rate. The authors also conclude that formaldehyde desorption is the rate-determining step.

Using the formaldehyde desorption rate at 600 K as the benchmark, the catalyst reduction rate is chosen to be a factor of 2 greater. The other reactions, which experimentally have been shown to be significantly faster, were chosen to be a factor of 10 greater. Because of the high uncertainty associated with these parameters, a nonlinear robustness analysis must be performed to examine the influence of large deviations from the values chosen.

The catalytic model by itself is incomplete. In the absence of the gas-phase reactions, formaldehyde is not oxidized and so its yield simply rises with residence time. Without the catalytic model, only trivial yields of formaldehyde are achievable. Therefore, both models are necessary. Simulation of the catalyst was done by treating the anionic vacancies and lattice oxygen sites as species and distributing them along the reactor.

4. Illustrative Computational Studies

The optimal control design formulation was applied to several test problems. In each example, an effective strategy for finding a cost functional minimum involved making a few runs with increasingly demanding objectives. The optimal flux profiles of the previous run were found to be good initial points for subsequent runs. The computer code employed for these simulations has also been used in previous applications[6,7,8,20] and has performed well. It is important to note that the present work serves to show the potential significance of optimally controlling the methanol conversion to formaldehyde rather than attempting to corroborate any specific reaction mechanism or establish an upper limit on the product yield.

The average iteration took about 29 min of CPU time on an R4000 IRIS Indigo. Although global optimality could not be guaranteed, it is evident that good quality solutions are obtained using the proposed algorithm.

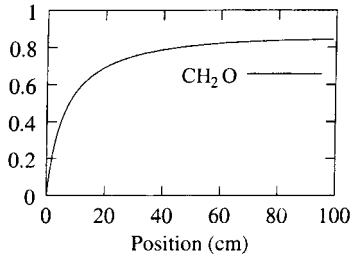

Figure 1. Mass fraction trajectory for formaldehyde from the 650 K isothermal reference case (1.45 s residence time).

In the examples, the length of the PFR in which the reactions proceed is $L = 100$ cm, with a cross sectional area of 40 cm^2. The optimal inlet flow rate at $l = 0$ is determined by trial and error. The reactor is at a constant pressure of 1 atm. The initial composition of the feedstock is chosen as 84% helium, 10% CH_3OH, and 6% O_2 by mass. Species mass fractions given in this paper are relative to the methanol co-fed into the reactor. In the simulations, the number of anionic vacancies and lattice oxygen sites is 200% of the moles of methanol feed so that it does not limit conversion.

The first simulation is the reference case with the base kinetic model of Table 1 under isothermal, 650 K conditions. The residence time is 1.45 s. The formaldehyde mass fraction at the reactor outlet is 0.84. The results are shown in Figure 1, and they are similar to those reported by Pernicone et al.[2] The goal of optimization is to improve upon the formaldehyde yield of this reference simulation.

A heat flux is the most important control variable, and the results of a heat flux optimization with the base kinetic model of Table 1 are shown in Figure 2. The optimized heat flux consists of two parts. The first part provides an initial pulse of heat to raise the temperature to the optimal zone for catalysis. In the second part, energy is then taken out to lower the temperature so as to slow the production of carbon monoxide and prevent formaldehyde yields from dropping. The heat flux optimization results in a significant drop in temperature in the second half of the reactor but only a modest increase in the first half. The final formaldehyde mass fraction has increased to 0.93, and the residence time is 1.13 s.

An oxygen flux does not increase methanol conversion to formaldehyde. This is consistent with results reported by Pernicone et al.[4] where varying oxygen composition did not effect methanol oxidation to formaldehyde over the same catalyst.

5. Robustness Analysis

The simulations of section 4 show that an optimization of heat flux with the base kinetic model resulted in an improved formaldehyde yield. The parameter values of the base kinetic model in Table 1 have considerable uncertainties. This calls for a thorough assessment of the robustness of the formaldehyde yield relative to the values chosen for the kinetic model parameters. This robustness assessment is carried out in two stages. First, section 5.1 explores the robustness of the formaldehyde yields relative to the values of the base kinetic model and also identifies the important parameters. This analysis is done in a fully nonlinear manner due to the large uncertainty ranges of the kinetic parameters.

The robustness analysis in section 5.1 is relative to a single optimal run with the base kinetic model, which represents the

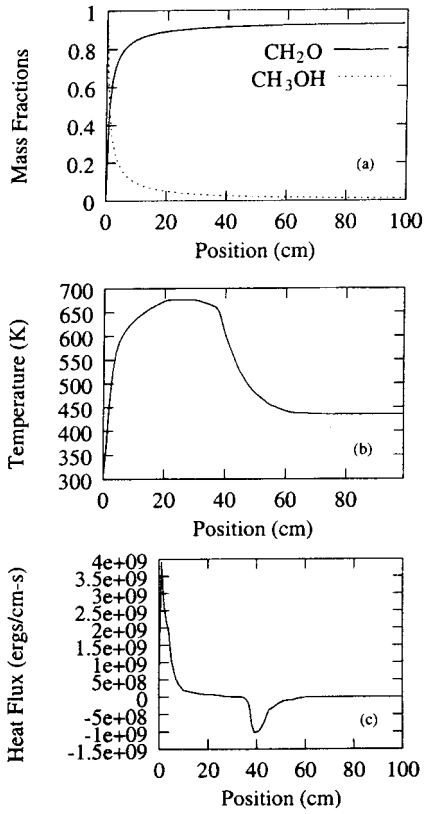

Figure 2. Optimization of heat flux with a 1.13 s residence time. (a) Mass fraction trajectories for formaldehyde and methyl alcohol. (b) Temperature profile of optimized solution. (c) Optimal heat flux.

best estimated values for the catalytic model parameters.[3,4] It was shown in section 4 that optimization of the heat flux improved the yield under the base kinetic model. It is important to show if optimization produces increased yields of formaldehyde for different models having other reasonable values for the kinetics parameters. If this is the case, a laboratory reactor with feedback optimization should produce an increase in formaldehyde yields regardless of the true values for the kinetic model parameters.

To address the latter global robustness issue, section 5.2 considers 100 kinetic systems. Each of these systems has a different set of random values for the kinetic model parameters. Each system is separately examined both under 650 K isothermal reference conditions and with optimization of the heat flux. The statistics of the reference and optimization behavior for the 100 runs gives very positive conclusions regarding the likelihood of an optimal distributed heat flux laboratory reactor producing a significant formaldehyde yield improvement.

5.1. Robustness to Model Uncertainties. The large uncertainties of the catalytic model do not lend themselves to conventional gradient sensitivity analysis. A new efficient fully nonlinear technique for robustness analysis may be performed by the random sampling-high dimensional model representation (RS-HDMR),[29] as presented in the Appendix. This technique efficiently represents the overall variance as a superposition of variances due to individual variables and groups of variables. The analysis is not applied to the gas-phase model, as it is far more well-defined than the catalytic model. Reasonable linear changes in the kinetic parameters of the gas-phase model resulted in no significant changes in the yield.

The catalytic model had nine input variables: k_1, k_2, k_3, k_5, A, E_a, a sticking coefficient, and the number of moles of anionic vacancies and lattice oxygen sites relative to the number of methanol moles fed into the reactor. The parameters A and E_a are components of $k_4 = A \exp[-E_a/RT]$. The following ranges were used in the robustness analysis: $10^9 \le k_1, k_3 \le 10^{12}$ L mol^{-1} s^{-1}, $10^9 \le k_5 \le 10^{12}$ atm$^{.5}$ s^{-1}, $10^7 \le k_2 \le 10^{11}$ L mol^{-1} s^{-1}, $10 \le E_a \le 20$ kcal, and $10^{12} \le A \le 10^{16}$ s^{-1}. The sticking coefficient was allowed to vary from 0 to 1. The number of moles of anionic vacancies and lattice oxygen sites relative to the number of methanol moles fed into the reactor was taken to vary from 0 to 2. The ranges are chosen to cover the data points of Liberti et al.[3] and to encompass large conservative deviations. The model output treated for robustness analysis is the mass fraction of formaldehyde in a simulation using the optimized heat profile shown in Figure 2. Legendre polynomials up to order 10 are used as the approximating bases for the expansion shown in eq B.7. A total of 1000 points are generated for the Monte Carlo integrals to calculate the RS-HDMR expansion up to second order. The convergence of the expansion was then tested against 1000 out-of-sample random points in the parameter space. The output of over 90% of the test points were predicted to an error of no more than 10% by the RS-HDMR approximation. The error is due to both the approximation of integrals using Monte Carlo integration and the truncation of the RS-HDMR expansion at second order. These tests confirm adequate convergence of the analysis.

TABLE 2: Largest Partial Variances σ_i^2 or σ_{ij}^2 and Global Robustness Indices[a]

variables	σ_i^2 or σ_{ij}^2	robustness index	variables	σ_i^2 or σ_{ij}^2	robustness index
LO	0.022	0.13	SC	0.0064	0.038
AV	0.020	0.12	k_2, SC	0.0055	0.033
E_a	0.014	0.095	A, LO	0.0047	0.028
A	0.012	0.081	k_3	0.0032	.019
k_2	0.011	0.068	E_a, LO	0.0025	0.015
A, E_a	0.010	0.062	AV, LO	0.0020	0.012
k_2, AV	0.0096	0.057	E_a, k_2	0.0018	0.011
k_1, SC	0.0091	0.054	E_a, AV	0.0016	0.0098
k_1	0.0066	0.039	A, AV	0.0015	0.0087

[a] SC is the sticking coefficient, LO is number of lattice oxygen moles relative to the number of methanol moles in the co-feed, and AV is the number of moles of anionic vacancies relative to the number of methanol moles in the co-feed.

The results of the RS-HDMR robustness analysis are presented in Table 2. The mean formaldehyde mass fraction is 0.71 compared to the optimal result of 0.93. The total variance is -0.219 and $+0.066$, as observed from the statistics of the 1000 Monte Carlo runs. Since yields are high, changes in the catalytic model parameter values that produce lower yields have a larger impact than changes that produce higher yields.

The role of each of the variables alone or cooperatively is captured by a set of robustness indices defined in eq B.18. The largest global robustness indices belong to the numbers of anionic vacancies and lattice oxygens. This implies that the most important variable is the available surface area of the catalyst. The next largest global robustness indices belong to E_a and A. This is consistent with the fact that the formaldehyde desorption is the rate-limiting step. The next slowest step is the reduction of the catalyst and thus k_2 also has a large robustness index. The sticking coefficient has only a modest robustness index and so it is reasonable to assume that unless the sticking coefficient is very low, it does not affect the kinetics to a great extent. The rate constants k_1, k_3, and k_5 have smaller senstivity indices, consistent with the fact that their respective reactions are fast.

5.2. Robustness of the Benefits of Optimization. The simulations in section 4 and the analysis of section 5.1 employed the base kinetic model with the best estimated values for the

parameters.[3,4] As these values have large uncertainties, it is important to show whether the resulting improvement in formaldehyde yields through heat flux optimization is robust to the values of the kinetic model parameters.

To explore this latter issue, 100 systems with random sets of values for the kinetic model parameters were taken over the window of variation defined in section 5.1. The optimization technique of section 4 was then applied to each of these systems. First, reference simulations were performed under 650 K isothermal conditions with no optimization. The mean formaldehyde mass fraction of these reference systems is 0.67 and the variance is 0.13. Then, each system was subjected to an optimal design of heat flux. The optimizations resulted in a mean increase in mass fraction of 0.073 with a variance of only 0.002. The small value of the variance around the mean mass fraction increase implies that each of the 100 widely different kinetic systems saw essentially the same 7% improvement in formaldehyde yield. This result suggests that employment of the reactor configuration in the laboratory with feedback optimization should produce a significant improvement in yield in spite of the lack of full quantitative understanding of the kinetic processes involved.

In summary, although the formaldehyde yields have significant uncertainty, the benefits of optimization are quite robust to the values assumed for the catalytic model parameters.

6. Conclusion

Optimal control of the conversion of methanol to formaldehyde in a PFR was considered in this paper. The chemical reaction model incorporated a $MoO_3-Fe_2(MoO_4)_3$ catalyst, and formaldehyde mass fractions of over 90% were achieved with an optimized heat flux. The improvement in yields from an optimized heat flux shows considerable robustness to the parameter values assumed in the catalytic model.

The components of the catalytic model were studied with a RS-HDMR robustness analysis technique that is more efficient than the ANOVA methodology in that it does not require a special sampling of the input for each partial variance. The most important variables were the catalyst surface area and the rate of formaldehyde desorption.

Under various conditions convergence to high-quality yields was achieved. Since global optimality was not guaranteed, better results are likely possible within the model. However, other nonmodeled physical and chemical processes could alter the yields in either direction.

The benefits of optimizing distributed fluxes shows that enhancement of the methanol conversion process can be achieved by employing special reactor configurations. In the laboratory, the solution designs can serve as a starting point for a reactor with feedback control. The output performance of the reactor will be fed to a learning algorithm, to in turn design the next experiment in a repeated sequence. This self-optimization is independent of any model assumptions and will therefore bring forth the true yields.

Acknowledgment. We acknowledge support from the National Science Foundation and the DOD.

Appendix

A. Optimal Control Formulation. The goal is to manipulate j and q, the input fluxes, so as to approach the optimal composition vector x^f while obeying the conservation equations described above. Satisfaction of dynamical constraints imposed by mass and energy conservation is assured by introducing Lagrange multipliers.

The objective function $J(x_i, j_i, q, T)$ for the optimal control problem can be formulated in many ways, but it is desirable that J be a smooth and convex function that is bounded from below.[22] The chosen objective function was

$$J = C + \sum_{i=1}^{n}\int_0^L \lambda_i\left[\frac{1}{F}(w_i - x_i\sum_{k=1}^{k=n} j_k + j_i) - \frac{\mathrm{d}x_i}{\mathrm{d}l}\right]\mathrm{d}l + \int_0^L \lambda_T\left[\frac{1}{F\sum C_{p_i}x_i}((T_0 - T)\sum_{i=1}^{n} C_{p_i}j_i - \sum_{i=1}^{n} H_{f_i}w_i - T\sum_{i=1}^{n} C_{p_i}w_i + q) - \frac{dT}{dl}\right]dl \quad \text{(A.1)}$$

where

$$C = \frac{1}{2}((x'(L) - x^f)^{\mathrm{T}}W_f(x'(L) - x^f)) + \frac{1}{2}\int_0^L x^{\mathrm{T}}W_x x\,\mathrm{d}l + \frac{1}{2}\int_0^L j^{\mathrm{T}}W_j j\,\mathrm{d}l + \frac{1}{2}\int_0^L qW_q q\,\mathrm{d}l \quad \text{(A.2)}$$

L is the total length of the reactor. C consists of four terms. The first term controls the composition of the mixture at the end of the reactor. The desired final composition is x^f and W_f is a positive weight matrix. The next three terms respectively minimize the concentrations of undesired species, species fluxes, and heat flux along the reactor. W_x and W_j are penalty weight matrices. W_q is a scalar penalty weight. The weight matrices are positive-definite and chosen to be diagonal to avoid any correlation terms. The equations describing the reactor are incorporated into J through the Lagrange multipliers λ_i and λ_T. Pontryagin's maximum principle[23] then allows us to find the sufficient conditions for optimality to be

$$\frac{\delta J}{\delta j_i} = \frac{\delta J}{\delta q} = \frac{\delta J}{\delta x_i} = \frac{\delta J}{\delta \lambda_i} = \frac{\delta J}{\delta \lambda T} = 0 \quad \text{(A.3)}$$

$$\lambda_i(L) = \frac{\partial J}{\partial x_i(L)} \quad \text{(A.4)}$$

$$\lambda T(L) = \frac{\partial J}{\partial T(L)} \quad \text{(A.5)}$$

The adjoint equations for the Lagrange multipliers are

$$\frac{\mathrm{d}\lambda_i}{\mathrm{d}l} = -W_{x_{ii}}x_i + \frac{1}{F}\left(-\sum_{k=1}^{n}\lambda_i\frac{\mathrm{d}w_k}{\mathrm{d}x_i} + \lambda_i\sum_{k=1}^{n} j_k + \frac{\lambda_T}{\sum C_{p_i}x_i}\left(\sum_{k=1}^{n} H_{fk}\frac{\mathrm{d}w_k}{\mathrm{d}x_i} + T\sum_{k=1}^{n} C_{p_k}\frac{\mathrm{d}w_k}{\mathrm{d}x_i}\right)\right) - \frac{\lambda_T C_{p_i}}{F(\sum C_{p_i}x_i)^2}((T_0 - T)\sum_{i=1}^{n} C_{p_i}j_i - \sum_{i=1}^{n} H_{f_i}w_i - T\sum_{i=1}^{n} C_{p_i}w_i + q) \quad \text{(A.6)}$$

$$\frac{\mathrm{d}\lambda_T}{\mathrm{d}l} = -\frac{1}{F}\sum_{i=1}^{n}\lambda_i\frac{\mathrm{d}w_i}{\mathrm{d}T} + \frac{\lambda_T}{F\sum C_{p_i}x_i}\left(\sum_{i=1}^{n}\left(H_{fi}\frac{\mathrm{d}w_i}{\mathrm{d}T} + \frac{\mathrm{d}H_{fi}}{\mathrm{d}T}w_i\right) + \sum_{i=1}^{n} C_{p_i}j_i + T\sum_{i=1}^{n} C_{p_i}\frac{\mathrm{d}w_i}{\mathrm{d}T} + \sum_{i=1}^{n} C_{p_i}w_i\right) \quad \text{(A.7)}$$

where $W_{x_{ii}}$ is the (i,i) element of W_x.

The end point conditions of the adjoint equations are found to be

$$\lambda_i(L) = W_{fi}(x_i'(L) - x^f) \quad \text{(A.8)}$$

$$\lambda_T(L) = 0 \tag{A.9}$$

where W_{fi} is the (i,i) entry in the W_f matrix.

The gradient of the objective function, which will be set equal to zero, can now be calculated:

$$\frac{\delta J}{\delta j_i} = W_{j_{ii}} j_i + \frac{\lambda_i}{F} - \frac{1}{F}\sum_{i=1}^{n}\lambda_i x_i - \int_l^L \frac{1}{F^2}\left(\sum_{i=1}^{n}\lambda_i(w_i - x_i\sum_{i=1}^{n} j_i + j_i) + \frac{\lambda_T}{\sum C_{p_i}x_i}((T_0 - T)\sum_{i=1}^{n} C_{p_i} j_i - \sum_{i=1}^{n} H_{f_i} w_i - T\sum_{i=1}^{n} C_{p_i} w_i + q)\right) dl + \frac{\lambda_T}{F\sum C_{p_i}x_i} C_{p_i}(T_0 - T) \tag{A.10}$$

$$\frac{\delta J}{\delta q} = W_q q + \frac{\lambda_T}{F\sum C_{p_i}x_i} \tag{A.11}$$

where $W_{j_{ii}}$ is the (i,i) element of W_j.

The following optimal control algorithm[20] for the two-point boundary value problem, based on the classical approach discussed in Hicks and Ray[24] and Jones and Finch,[25] was used:

1. Decide on objective function and choice of influxes.

2. Discretize reactor into N points l_k, $k = 1 \ldots N$. Assign initial values for $j_i(l_k)$ and $q(l_k)$.

3. Apply cubic spline to $j_i(l_k)$ and $q(l_k)$ to get a continuous representation of j_i and q. Knowing $x_i(0)$, $j(l)$, and $q(l)$, solve equations of motion to obtain $x(l)$ and $T(l)$.

4. Knowing $x(l)$, solve for adjoint initial conditions and then integrate adjoint equations to get λ_i's and λ_T.

5. Knowing $x(l)$, $T(l)$, and λ's, evaluate gradient at each discretization point.

6. If the norm of the gradient calculated in the previous step is less than a tolerance parameter, terminate algorithm. Otherwise, use gradient in the conjugate gradient minimizer to update the values of $j_i(l_k)$ and $q(l_k)$ and go back to step 3.

7. Use sensitivity analysis to determine which reactions were of greatest significance to the solution obtained.

In the calculations, the CONMIN[26] code was used as the conjugate gradient minimizer; the chemical kinetics package CHEMKIN-II[27] was employed to interface the thermodynamic and kinetics data, and LSODA[28] was used as a differential equation integrator.

To ensure positive mass flux densities j_i, a transformation to a new control variable j_i' was done where $j_i = \exp(j_i')$. The gradient is then modified to

$$\frac{\delta J(l)}{\delta j_i'(l)} = \exp(j_i')\frac{\delta J}{\delta j_i(l)} \tag{A.12}$$

B. High Dimensional Model Representation Formulation. The high dimensional model representation (HDMR) technique captures the input−output relationships of physical systems with many input variables.[29,30] The HDMR assumes a function can be represented as the following exact hierarchical expansion:

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{ij}(x_i,x_j) + \ldots + f_{12\ldots n}(x_1,x_2,\ldots,x_n) \tag{B.1}$$

Here f_0 is a constant, $f_i(x_i)$ represents the effect of x_i upon the output, and $f_{ij}(x_i, x_j)$ represents the cooperatve effect of the variables x_i and x_j. The effect of increasing numbers of cooperating variables acting together is represented by higher order terms. Experience shows that for many physical systems, only terms up to second order need be considered.

The analysis of variance (ANOVA) decomposition is used in statistics to analyze variances by representing multivariate functions as a superposition of functions of fewer variables.[31,32] The component functions $f_{i_1\ldots i_p}(x_{i_1}, \ldots, x_{i_p})$ are required to be orthogonal, and this property is assured by the relation

$$\int_{[0.1]} f_{i_1\ldots i_p}(x_{i_1},\ldots,x_{i_p})\, dx_k = 0 \qquad \text{for } k = i_1, i_2, \ldots, i_p \text{ and } p = 1, 2, \ldots, l \tag{B.2}$$

where [0,1] denotes the integration interval.

The component functions of the ANOVA expansion may then be calculated using the following integrals:

$$f_0 \equiv \int_{[0,1]^n} f(x)\, dx \tag{B.3}$$

$$f_i(x_i) \equiv \int_{[0,1]^{n-1}} f(x)\prod_{j\neq i} dx_j - f_0 \tag{B.4}$$

$$f_{ij}(x_i, x_j) \equiv \int_{[0,1]^{n-2}} f(x)\prod_{k\notin\{i,j\}} dx_k - f_i(x_i) - f_j(x_j) - f_0 \tag{B.5}$$

$$f_{i_1\ldots i_l}(x_{i_1}, .., x_{i_l}) \equiv \int_{[0,1]^{n-l}} \prod_{k\notin\{i_1,\ldots,i_l\}} dx_k - \sum_{j_1<\ldots<j_{l-1}\subset\{i_1,\ldots i_l\}} f_{j_1,\ldots,j_{l-1}} - \sum_{j_1<\ldots<j_{l-2}\subset\{i_1,\ldots i_l\}} - f_{j_1,\ldots,j_{l-2}} - \ldots - \sum_j f_j(x_j) - f_0 \tag{B.6}$$

These integrals are determined by Monte Carlo Integration.[33] The evaluation of these Monte Carlo approximations requires sampling the input on a regular net, an expensive, if at all feasible, procedure.[33]

A more efficient random sampling-high dimensional model representation (RS-HDMR) that avoids the latter problem can be developed.[30] The following expansion is used to represent a multivariate function:

$$f(x) = c_0 + \sum_{i=1}^{n}\sum_{j=1}^{s} c_{ij}\phi_{ij}(x_i) + \sum_{i<j}\sum_{k=1}^{s} c_{ijk}\phi_{ijk}(x_i,x_j) + \ldots \tag{B.7}$$

$\{\phi_i(x_j)\}_{i=1}^{s}$ is a family of approximating bases for the univariate functions of the variable x_j on the unit interval [0, 1]. $\{\phi_{ijk}(x_i, x_j)\}_{k=1}^{s}$ is the approximating family for the bivariate functions. All these functions have a zero mean to satisfy eq B.1, but the functions do not have to be orthogonal.

The coefficients c_0, $\{c_{ij}\}$, and $\{c_{ijk}\}$ are calculated with the following integrals using the properties of the basis functions defined above:

$$c_0 = \int_{[0,1]^n} f(x)\, dx \tag{B.8}$$

$$M^i\begin{bmatrix} c_{i1} \\ \vdots \\ c_{is} \end{bmatrix} = \begin{bmatrix} \int_{[0,1]^n} f(x)\,\phi_{i1}(x_i)\, dx \\ \vdots \\ \int_{[0,1]^n} f(x)\,\phi_{is}(x_i)\, dx \end{bmatrix} \qquad i = 1, \ldots, n \tag{B.9}$$

$$M^{ij}\begin{bmatrix} c_{ij1} \\ \vdots \\ c_{ijs} \end{bmatrix} = \begin{bmatrix} \int_{[0,1]^n} f(x)\,\phi_{ij1}(x_i,x_j)\,\mathrm{d}x \\ \vdots \\ \int_{[0,1]^n} f(x)\,\phi_{ijs}(x_i,x_j)\,\mathrm{d}x \end{bmatrix} \qquad j = 1, ..., n;\ i < j \tag{B.10}$$

where the matrix families $\{M^i\}$ and $\{M^{ij}\}$ have the elements:

$$M^i_{kl} = \int_{[0,1]} \phi_{ik}(x_i)\,\phi_{il}(x_i)\,\mathrm{d}x_i \tag{B.11}$$

$$M^{ij}_{kl} = \int_{[0,1]^2} \phi_{ijk}(x_i,x_j)\,\phi_{ijl}(x_i,x_j)\,\mathrm{d}x_i\,\mathrm{d}x_j \tag{B.12}$$

If the basis functions used are orthogonal, the following Monte Carlo approximations may be used:

$$c_0 = \int_{[0,1]^n} f(x)\,\mathrm{d}x \approx \frac{1}{N}\sum_{r=1}^{N} f(x^r_1,x^r_2,...,x^r_n) \tag{B.13}$$

$$c_{ij} = \int_{[0,1]^n} f(x)\,\phi_{ij}(x_i)\,\mathrm{d}x \approx \frac{1}{N}\sum_{r=1}^{N} f(x^r_1,x^r_2,...,x^r_n)\,\phi_{ij}(x^r_i) \tag{B.14}$$

$$c_{ijk} = \int_{[0,1]^n} f(x)\,\phi_{ijk}(x_i, x_j)\,\mathrm{d}x \approx \frac{1}{N}\sum_{r=1}^{N} f(x^r_1,x^r_2,...,x^r_n)\,\phi_{ijk}(x^r_i,x^r_j) \tag{B.15}$$

where N is the sample size. The error is of the order $(1/\sqrt{N})$ and is weakly dependent on the dimension. Thus, unlike the ANOVA integrals, no costly sampling is required and the number of model runs needed is invariant with the number of parameters.

For random inputs consisting of independently distributed variables, the component functions are uncorrelated and so the overall variance can be written as a superposition of variances due to individual variables and groups of variables:

$$\sigma = \mathbf{E}(f - f_0)^2 = \sum_i \sigma_i + \sum_{i<j} \sigma_{ij} + ... \sigma_{12....n} \tag{B.16}$$

where the individual variances are given by

$$\sigma_{i_1,i_2,...,i_l} = \int_{K^l} (f_{i_1...i_l})^2\,\mathrm{d}x_{i_1...dxi_l} \tag{B.17}$$

The global robustness indices[27,28,29] are defined as

$$S_{i_1,i_2,...,i_l} = \frac{\sigma_{i_1,i_2,...,i_l}}{\sigma} \tag{B.18}$$

These indices represent the fractional contribution of the input set $\{x_{i_1}, ... x_{i_l}\}$ to the variance of the output. Because they do not rely on small pertubations, these indices may be used for nonlinear robustness analyses where the input parameters are characterized by large uncertainty.

References and Notes

(1) Held, T. J.; Dryer, F. L. A Comprehensive Mechanism for Methanol Oxidation. Submitted to *Int. J. Chem. Kinet*.
(2) Pernicone, N. *J. Less-Common Met.* **1974**, *36*, 289.
(3) Liberti, G.; Pernicone, N.; Soattini, S. *J. Catal.* **1972** *27*, 52.
(4) Pernicone, N.; Lazzerin, F.; Liberti, G.; Lanzavecchia, G. *J. Catal.* **1969**, *14*, 293.
(5) Batist, PH. A.; Lippens, B. C.; Schuit, G. C. A. *J. Catal.* **1966**, *5*, 55.
(6) Faliks, A.; Yetter, R. A.; Floudas, C. A.; Hall, R.; Rabitz, H. *J. Phys. Chem. A* **2000**, *104*, 10740.
(7) Faliks, A.; Yetter, R. A.; Floudas, C. A.; Hall, R.; Rabitz, H. Submitted to *Ind. Eng. Chem. Res.*
(8) Faliks, A.; Yetter, R. A.; Floudas, C. A.; Wei, Y.; Rabitz, H. *Polymer* **2001**, *42*, 2061.
(9) Dryer, F. L. *Fossil Fuel Combustion: A Sourcebook*; Bartok, W., Sarofim, A. F., Eds.; Wiley: New York, 1991.
(10) Yetter, R. A.; Dryer, F. L.; Rabitz, H. *Combust. Sci. Technol.* **1991**, *79*, 97.
(11) Hochgreb, S.; Dryer, F. L. *Combust. Flame* **1992**, *91*, 257.
(12) Held, T. J.; Dryer, F. L. *25th Symp. (Int.) Combust.* **1994**, 901.
(13) Fame, G.; Pernicone, N. *Rend. Accad. Naz. Lincei.* **1968**, VIII, *44*, 560.
(14) Gorokhovatski, Y. B.; Evmenenko, N. P. *Kinet. Catal.* **1969**, *10*, 1071.
(15) Gorokhovatski, Y. B.; Evmenenko, N. P. *Kinet. Catal.* **1970**, *11*, 104.
(16) Bibin, V. N.; Popov, B. I. *Kinet. Catal. (USSR)* **1969**, *10*, 1091.
(17) Popov, B. I.; Osipova, K. D. *Kinet. Catal. (USSR)* **1971**, *12*, 642.
(18) Popov, B. I.; Osipova, K. D.; Pankratev, Y. D. *Kinet. Catal. (USSR)* **1971**, *12*, 1102.
(19) Popov, B. I.; Osipova, K. D.; Malakhov, V. F.; Kolchin, A. M. *Kinet. Catal. (USSR)* **1971**, *12*, 1297.
(20) Rojnuckarin, A.; Floudas, C. A.; Rabitz, H.; Yetter, R. *Ind. Eng. Chem. Res.* **1996**, *35*, 683.
(21) Kramer, M. A.; Calo, J. M.; Rabitz, H. *Appl. Math. Model* **1981**, *5*, 432.
(22) Leigh, J. R. *Functional Analysis and Control Theory*; Academic Press: New York, 1980.
(23) Bryson, A. E.; Ho., Y. *Applied Optimal Control*; Blaisdell Publishing: New York, 1973.
(24) Hicks, G. A.; Ray, W. H. *Can. J. Chem. Eng.* **1971**, *49*, 522.
(25) Jones, D. I.; Finch, J. W. *Int. J. Control* **1984**, *40*, 747.
(26) Shanno, D. F.; Phua, K. H. Minimization of unconstrained multivariable functions. *ACM Transactions on Mathematical Software* **1980**, *6*, 618.
(27) Kee, R. J.; Rupley, F. M.; Miller, J. A. *CHEMKIN-II A Fortran Chemical Kinetics Package*. Sandia National Laboratories, 1989.
(28) Hindmarsh, A. C. *Odepack, A Systemized Collection of Ode Solvers in Scientific Computing*. North-Holland, 1983.
(29) Rabitz, H.; Alis, O. F. *J. Math. Chem.* **1999**, *25*, 197.
(30) Alis, O. F.; Rabitz, H. *J. Math. Chem.*, in press.
(31) John, P. W. M. *Statistical Design and Analysis of Experiments*; MacMillian Press: New York, 1971.
(32) Edwards, A. L. *Multiple Regression and the Analysis of Variance and Covariance*; W. H. Freeman: New York, 1985.
(33) Shreider, Y. *The Monte Carlo Methodol*; Pergamon Press: Oxford, U.K., 1967.
(34) Sobol, I. *USSR Computational Math., Math. Phys.* **1976**, *7*, 86.
(35) Saltelli, A.; Sobol, I. *Reliability Engineering and System Safety* **1995**, *50*, 225.
(36) Homma, T.; Saltelli, A. *Reliability Engineering and System Safety* **1996**, *52*, 1.