AIChE

Optimal Design of Nonlinear Temperature Programmed Reduction Experiments

Peter Heidebrecht
Physical and Chemical Process Engineering, Max Planck Institute, Sandtorstrasse 1, 39106 Magdeburg, Germany

Kai Sundmacher
Physical and Chemical Process Engineering, Max Planck Institute, Sandtorstrasse 1, 39106 Magdeburg, Germany

Lorenz T. Biegler
Dept. Chemical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213

DOI 10.1002/aic.12485
Published online January 19, 2011 in Wiley Online Library (wileyonlinelibrary.com).

We propose the application of nonconstant temperature gradients to improve the quality of temperature programmed reduction (TPR) experiments with respect to parameter estimation and model discrimination. This leads to TPR experiments with nonlinear temperature profiles (N-TPR). To determine optimal profiles for the temperature gradient, optimal control problems are set up and solved numerically. The results show that N-TPR experiments can be significantly better than traditional linear TPR experiments for many different scenarios. To implement these results in practice, we develop and demonstrate reduced optimization problem formulations, which can be solved faster and more reliably than the original formulation, with very similar results.
 AIChE J, 57: 2888–2901, 2011
Keywords: temperature programmed reduction, experimental design, optimal control, model reduction

Introduction

Temperature Programmed Reduction (TPR) is an experimental method that is widely used for the estimation of kinetic parameters of metal oxide reduction reactions.[1,2] Its working principle is simple (Figure 1); hydrogen (or some other reducing gas) flows continuously through a small sample of the metal oxide powder. The metal oxide is reduced in the hydrogen atmosphere, converting part of the hydrogen to steam. The concentration of steam in the exhaust gas is proportional to the reaction rate in the sample and is measured in short time intervals (about one measurement per second). The reduction process may include several reaction steps in which the steam fraction is proportional to the cumulated rates of all reactions. The sample temperature starts at low values (ambient temperature up to 500 K), is increased at a constant rate, and follows a linear profile over time. Typical temperature gradients are between 5 and 25 K/min. At initial time, due to the low temperatures, reaction rates are virtually zero. With increasing temperature, the reaction rates increase and eventually, reactions reach complete conversion. As soon as full reduction conversion is reached, the reaction rates approach zero again, and the TPR run ends. The resulting measurement signal is used to estimate kinetic parameters. It may also be applied to discriminate rivaling models.[3–5]

The only control variable that can be changed from one TPR experiment to another is the applied temperature

Correspondence concerning this article should be addressed to P. Heidebrecht at heidebrecht@mpi-magdeburg.mpg.de.

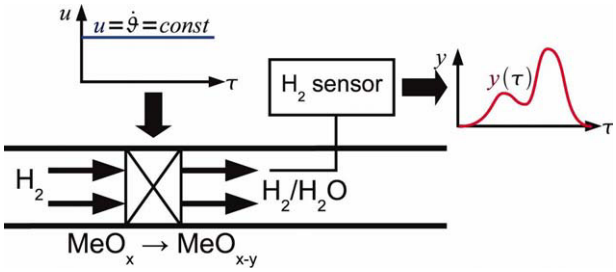

Figure 1. Principle of a TPR experiment.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

gradient. Typically, a series of three to five TPR experiments is carried out with different temperature gradients and the whole ensemble of measured profiles is used for the parameter estimation. In a system with three reactions [e.g., the reduction of hematite (Fe_2O_3)], 12 parameters need to be estimated from these few experiments: the reaction rate constants, the activation energies, the orders of reaction, and the fractions of convertible material in the sample.

With 12 unknown parameters and only one single control parameter, the TPR method has very limited options to control the experiment. This can also be seen in a contribution by Pineau et al.,[6] where they present a compilation of TPR studies on the reaction kinetics of iron oxide reduction carried out by various authors. Essentially, they show that estimates of the activation energies of reduction reactions vary significantly, even though they were obtained under comparable conditions. While some of these differences may be attributed to morphological differences in the sample materials, this result indicates that the parameter estimates have a large uncertainty. Moreover, with regard to model discrimination, we have shown[3] that the classical TPR method is not well suited to identify the reaction scheme or the type of reaction kinetics of systems with multiple reactions; the data from a series of TPR experiments could be fitted equally well by very different models.

One option to amend these deficits of the TPR method is to extend its control options. This can be achieved by lifting the restriction of a constant temperature gradient and thereby turning it into a function of time. Because this leads to experiments with nonlinear temperature profiles, we propose to call this method N-TPR (Nonlinear Temperature Programmed Reduction). The control function, namely the temperature gradient, can then be designed in such a way that the resulting measurement function becomes optimal for the goals of the experiment. This leads to formulation of optimal control problems, which are developed and described in Section "Problem Formulation." These are solved numerically, and some exemplary solutions are shown in Section "Optimal N-TPR Experiments." With respect to better applicability of N-TPR, we propose to use reduced optimization problems, which can be solved reliably and quickly. They are introduced and discussed in Section "Reduced Problem Formulation."

Problem Formulation

In this study, we develop optimization problem formulations for two different purposes: to design optimal N-TPR experiments for parameter estimation (Section "Optimal control problem for improved parameter estimation") and to discriminate among competing models (Section "Optimal control problem for model discrimination"). Because the dynamic model equations of a TPR system are common to both problems, we introduce them in the next section.

The TPR model

The TPR model describes the relation between the input and the output of a TPR experiment (Figure 2). The only input variable is the temperature gradient. In traditional TPR, this is a constant, while in N-TPR, it is time dependent. The output variable is the measurement signal that essentially corresponds to the cumulative reaction rate in the sample. The model states comprise the sample composition, x, and temperature, ϑ, which we normalize by a fixed standard temperature.

The model is formulated in terms of the following dimensionless parameters, p (see Heidebrecht et al.[3] for a more detailed derivation):

- The Damköhler number, Da_j, corresponds to the reaction rate constant of reaction j at reference temperature, $\vartheta_j^{\mathrm{ref}}$.
- The Arrhenius number, $\mathrm{Arr}_j = E_j/\mathrm{RT}^0$, corresponds to the activation energy of reaction j.
- The oxygen capacity, Θ_j, corresponds to the fraction of oxygen that is released by reaction j, related to the total amount of reducible oxygen in the fully oxidized material.
- The order of the reaction, n.

In addition, we define the rate of reaction j, that is, $R_j(x, \vartheta, p)$ as well as the temperature gradient, u. The model equations include the mass balance, which comprises the reaction rates, $\mathrm{Da}_j \cdot R_j$, divided by the oxygen capacity, Θ_j and multiplied by the stoichiometric coefficients, $\nu_{i,j}$. This balance describes the change in the sample composition, x:

$$\dot{x}_i = f(x, \vartheta, p) = \sum_{j=1}^{N_r} \nu_{i,j} \cdot \frac{\mathrm{Da}_j}{\Theta_j} \cdot R_j(x, \vartheta, p); \quad i = 1 \ldots N_s \tag{1}$$

with N_s denoting the number of reducible species and N_r the number of reactions in the system. The initial conditions for these ODEs are

$$x(\tau = 0) = x_0 \tag{2}$$

The temperature is increased according to the temperature gradient:

$$\dot{\vartheta} = u(\tau) \tag{3}$$

with the initial temperature

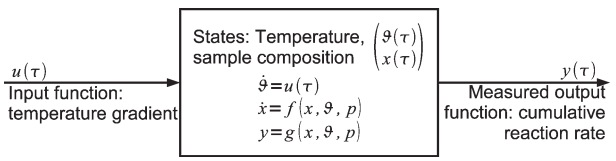

Figure 2. Input-output scheme of a TPR experiment.

$$\vartheta(\tau = 0) = \vartheta_0 \tag{4}$$

The measured output signal is the sum of all reaction rates:

$$y = g(x, \vartheta, p) = \sum_{j=1}^{N_r} \mathrm{Da}_j \cdot R_j(x, \vartheta, p) \tag{5}$$

The reaction kinetics comprise an Arrhenius term, considering the effect of temperature on the reaction rate, as well as a composition dependent function. The latter depends on the mole fraction of the reactant species of reaction j, $x_{\tilde{j}}$, and has the reaction order n_j.

$$R_j = \exp\left(\mathrm{Arr}_j \cdot \left(\frac{1}{\vartheta_j^{\mathrm{ref}}} - \frac{1}{\vartheta}\right)\right) \cdot r_j\left(x_{\tilde{j}}, n_j\right); \quad j = 1 \ldots N_r \tag{6}$$

In solid state reactions, three kinetic laws are commonly applied: Power law, Avramy-Erofeev, and diffusion-limited kinetics.[7,8] Additional kinetic models such as the shrinking core or the crackling core models[9] may be applied, but are not considered here.

- *Power law kinetics*

$$r_j = x_{\tilde{j}}^{n_j}; \quad 0.5 \le n_j \le 2.0 \tag{7}$$

Strictly speaking, the physical motivation behind the power law kinetics—the probability of n_j reactant molecules meeting at a reaction site at the same time—does not apply to solid phase conversions.[4,10] However, due to its simplicity and effectiveness in practical applications, power law kinetics are frequently used to describe reduction processes. The bounds on the reaction order in Eq. 7 are arbitrary, but they reflect a typical span for these kinetics.

- *Avramy-Erofeev kinetics*

$$r_j = \frac{x_{\tilde{j}} \cdot \left[-\log\left(x_{\tilde{j}}\right)\right]^{n_j}}{1 - n_j}; \quad \frac{1}{2} \le n_j \le \frac{3}{4} \tag{8}$$

The Avrami-Erofeev kinetics result from a simplification of a rather complex model that considers populations of nucleates of product species, each of a different size, in combination with a growth rate at the boundaries of nucleates, where product and reactant phase are in contact with each other. Equation 8 is derived for a single reaction system, with pure reactants as initial conditions.[11,12]

- *Diffusion limited kinetics*

$$r_j = \frac{x_{\tilde{j}}^{1-1/n_j}}{1 - x_{\tilde{j}}^{1/n_j}}; \quad 1 \le n_j \le 3 \tag{9}$$

Similar to Avrami-Erofeev kinetics, diffusion limited kinetics are derived with the assumption of pure reactants at initial time,[9] so its physical motivation is only valid for a single reaction system. At initial time, the reaction rate (Eq. 9) according to this formulation reaches infinite values. The overall rate (Eq. 6) remains bounded by choosing a zero initial temperature, and, consequently, zero Arrhenius term at initial time.

Optimal control problem for improved parameter estimation

Objective function We first consider N-TPR optimization problems that maximize the quality of the parameter estimates. The idea is to determine a control profile [the temperature gradient, $u(\tau)$], which results in a measurement signal [the cumulative reaction rate, $y(\tau)$] that offers the best covariance matrix of the estimated parameters. This matrix includes variances of individual parameters, as well as the pairwise correlation of different parameters, representing the interdependence of these estimates.

Several optimality criteria have been derived in an effort to define the "best" covariance matrix[13,14]:

- A-optimality criterion minimizes the trace of the covariance matrix. Thus, the individual estimates of the parameters are improved, but the off-diagonal elements describing the covariance between different parameters are ignored.
- C-optimality criterion minimizes the variance of an estimator of a linear function of the parameters. As with the A-optimality criterion, off-diagonal elements are not considered.
- D-optimality criterion minimizes the determinant of the covariance matrix. Thus, the interdependence of the estimated parameters is considered in this criterion.
- E-optimality criterion minimizes the largest eigenvalue of the covariance matrix. All other properties of the covariance matrix are disregarded. This criterion corresponds to minimizing the largest diameter of the parameter confidence region. The objective function may not be continuously differentiable at all points.

The D-optimality criterion considers the covariance elements of the matrix, reducing variances, and covariances alike, and is a continuously differentiable function. Moreover, by modifying the original D-optimality criterion, one can also choose to design experiments that minimize the variances of specific parameters or reduce the covariance of a certain pair of parameters. Because this criterion offers greater flexibility, it is the criterion that we consider in this study.

The objective function according to the D-optimality criterion is given by (see Bard,[15] Chapter 10):

$$\max \det\left(V_0^{-1} + B^T \Pi^{-1} B\right) \tag{10}$$

In this function, V_0 denotes the *a priori* covariance matrix of the parameters from previous estimates, Π denotes the expected covariance matrix of the measurements in the planned experiment, and B describes the expected sensitivity of the measurements with respect to the parameters in the planned experiment.

In the case of N-TPR, this objective function can be further modified according to the following assumptions. First, we note that *prior* parameter values as well as the covariance matrix, V_0, have to be estimated from previous experiments, which may be traditional linear or nonlinear TPR experiments. One can expect that *prior* parameter values of the model have a strong impact on the optimal design.

However, if V_0 is unknown it requires many optimizations to be conducted to cover a wide space of possible parameter values. Including the dependence of the optimal experiments on the a priori covariance matrix would increase the multitude of the necessary calculations. Therefore, to keep the number of cases to be solved to a reasonable level, we omit this matrix from the objective function. Of course, if V_0 is known *a priori*, it can easily be included in the objective function.

Second, the covariance matrix of the measurements, Π, is difficult to obtain because (a) the measurement error is not independent of the measured value and (b) the measurements in a single TPR run are not statistically independent. Thus, diagonal elements are of different magnitudes and off-diagonal elements in Π are non-zero. Because this is the *a priori* covariance matrix of the measurements, one would need a good model to predict it for TPR experiment, which is not available. Instead, for the purpose of planning the experiments, we replace the unknown matrix Π by an identity matrix. This essentially ignores the statistical properties of the measurements and, instead, focuses on the sensitivities from the postulated model. With that, we end up with a reduced objective function:

$$\max \det\left(B^T B\right) \tag{11}$$

This essentially tells us to design the experiment in such a way that its outcome is as sensitive with respect to the model parameters as possible.

The sensitivity matrix, B, comprises the sensitivities of the measurement signal with respect to the estimated parameters, p, at the designed experimental conditions (Bard,[15] Chapter 7–51). To account for different orders of magnitude of the parameters, the sensitivities are normalized by the actual value of the corresponding parameter. Let n be the number of measurements taken during a TPR run, τ_i the time when the i-th measurement is taken, and m the number of model parameters to be estimated from the TPR experiment. Then, the sensitivity matrix B can be estimated using the parameter sensitivities of the model output at each measurement point:

$$B \approx \begin{pmatrix} \left.\frac{\partial y}{\partial p_1}\right|_{\tau_1} \cdot p_1 & \cdots & \left.\frac{\partial y}{\partial p_m}\right|_{\tau_1} \cdot p_m \\ \vdots & & \vdots \\ \left.\frac{\partial y}{\partial p_1}\right|_{\tau_n} \cdot p_1 & \cdots & \left.\frac{\partial y}{\partial p_m}\right|_{\tau_n} \cdot p_m \end{pmatrix} = \begin{pmatrix} y_{p_1}(\tau_1) & \cdots & y_{p_m}(\tau_1) \\ \vdots & & \vdots \\ y_{p_1}(\tau_n) & \cdots & y_{p_m}(\tau_n) \end{pmatrix} \tag{12}$$

With this, the objective function reads:

$$\max \det B^T B \approx \begin{pmatrix} \sum_i y_{p_1}(\tau_i) \cdot y_{p_1}(\tau_i) & \cdots & \sum_i y_{p_1}(\tau_i) \cdot y_{p_m}(\tau_i) \\ \vdots & & \vdots \\ \sum_i y_{p_m}(\tau_i) \cdot y_{p_1}(\tau_i) & \cdots & \sum_i y_{p_m}(\tau_i) \cdot y_{p_m}(\tau_i) \end{pmatrix} \tag{13}$$

To adapt to continuous measurement signals, we shift from discrete, equidistant points in time to a continuous formulation. Thus, the summations are replaced by integrals over the whole duration, $(\tau = 0 \cdots \tau_e)$, of the TPR run:

$$\max \det \int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau \tag{14}$$

where

$$\begin{aligned} p &= (p_1, \ldots, p_m)^T \\ &= (\mathrm{Da}_1, \mathrm{Arr}_1, n_1, \Theta_1 \ldots \mathrm{Da}_{Nr}, \mathrm{Arr}_{Nr}, n_{Nr}, \Theta_{Nr})^T \end{aligned}$$

Optimal control problem The time dependent sensitivity functions, $y_{\mathrm{p}}(\tau)$, are obtained from direct sensitivity equations, the Jacobian ODEs, which pose additional constraints to the optimization problem. In addition to these ordinary differential equations and algebraic equations (DAEs), several inequality constraints are also imposed. The resulting optimization problem is given below, with N_{s} as the number of reducible species in the reaction system and N_{p} as the number of model parameters:

$$\max_{u(\tau)} \det \int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau \quad \text{Objective function} \tag{15}$$

$$\dot{x} = f(x, \vartheta, p) \quad \text{ODE constraint, sample composition} \quad \in \Re^{N_s} \tag{16}$$

$$\dot{\vartheta} = u(\tau) \quad \text{ODE constraint, temperature} \quad \in \Re^1 \tag{17}$$

$$\dot{J} = f_x \cdot J + f_p \quad \text{ODE constraint, Jacobian} \quad \in \Re^{N_s \times N_p} \tag{18}$$

$$y = g(x, \vartheta, p) \quad \text{AE constraint, output signal} \quad \in \Re^1 \tag{19}$$

$$y_p = \left(g_x \cdot J + g_p\right) \cdot p \quad \text{AE constraints, sensitivities} \quad \in \Re^{N_p} \tag{20}$$

$$x(\tau = 0) = x_0 \quad \text{Initial condition, Eq.16} \quad \in \Re^{N_s} \tag{21}$$

$$\vartheta(\tau = 0) = \vartheta_0 \quad \text{Initial condition, Eq. 17} \quad \in \Re^1 \tag{22}$$

$$J(\tau = 0) = 0 \quad \text{Initial condition, Eq. 18} \quad \in \Re^{N_s \times N_p} \tag{23}$$

$$u_{\min} \le u \le u_{\max} \quad \text{Bounds on control variable} \quad \in \Re^1 \tag{24}$$

$$\vartheta_{\min} \le \vartheta \le \vartheta_{\max} \quad \text{Bounds on temperature} \quad \in \Re^1 \tag{25}$$

$$x(\tau_e) \le x_{e,\max} \quad \text{Complete conversion at end time} \quad \in \Re^{N_s} \tag{26}$$

With regard to real experimental systems, which cannot realize arbitrarily high temperature gradients, an upper bound and a lower bound is imposed on the control variable in Eq. 24. The temperature itself is also limited. TPR devices usually do not have a cooling device, so the lowest applicable temperature is ambient temperature. Due to limited material stability, an upper bound is also imposed in Eq. 25. To ensure that the reduction has reached virtually complete conversion after the given duration of the run, τ_{e}, additional

lower bounds are introduced in Eq. 26 with a typical value of $x_{e,max} = 10^{-3}$.

For a system with three reactions, where 12 parameters have to be estimated, this optimal control problem has 40 ODEs and 13 AEs. Because of the Arrhenius terms, the state equations are nonlinear. The control profile, however, appears linearly in (Eq. 17). This leads to a so-called singular optimal control problem, which poses a number of challenges that will be explored in Section "Reduced Problem Formulation." To handle these problems, we first describe the following solution strategy.

Numerical treatment To solve the optimal control problem (Eqs. 15–26), the DAEs are discretized in time according to the method of orthogonal collocation on finite elements (see, e.g., Biegler[16]). Here, three collocation points are chosen per finite element, and the number of finite elements is varied between 100 and 300. To reduce the number of degrees of freedom, the optimization variable, $u(\tau)$ is assumed constant over each finite element. With that, a typical optimization problem for an N-TPR has about 10,000 variables and 100 degrees of freedom. The discretized problem is implemented in AMPL[17] and solved using the optimization algorithms CONOPT and IPOPT.[18]

Only for some parameter configurations, the optimization algorithm converges towards an optimum from any arbitrary initial point. Thus, the following procedure is proposed to enhance convergence:

- Initialization: The constraint equations are solved for some constant temperature gradient, that is, with no degrees of freedom. This is achieved by setting identical lower and upper bounds for the temperature gradients.
- Define the objective function and impose an inequality constraint for complete conversion (Eq. 26).
- Specify the upper and lower bounds on the temperature gradient in a sequence of relaxations. Each time one or both bounds are changed, the problem is solved again. This is a "trial and error" procedure and the algorithms IPOPT and CONOPT tend to work well in different cases, although performance of each is hard to predict in advance. If both algorithms fail, the whole procedure is started again, but with a more conservative relaxation of the bounds.
- An optimization run is usually performed for a given duration of the TPR experiment. Starting from an initial profile that stretches the TPR curve over the whole time span may be advantageous. Cases have been observed where starting from a profile with a maximum temperature gradient (and thus a very short TPR curve) do not converge to an optimum, but starting from a lower temperature gradient (with a TPR curve using almost the whole given time span) would end up in an optimum.
- Using up to six relaxation steps, the bounds reach the desired values and an optimal solution is obtained.

Typical solution times on a standard PC (Intel Core2 Duo CPU E6850, 3.00 GHz) is between 5 and 30 CPU minutes, depending on the number of relaxation steps required to achieve convergence.

Application strategy We expect N-TPR experiments to be applied in a campaign to estimate kinetic parameters for a given reaction system. Initially, when no estimates of the parameters are available, a few linear TPR experiments are conducted, so that a first estimate of the parameters can be obtained, together with a covariance matrix (V_0) of these estimates. If the covariances are larger than desired, then an N-TPR experiment is designed from the optimal control problem, based on these parameter estimates and their covariance. After conducting this experiment, new estimates are obtained from the data of the linear and the nonlinear experiments and the covariance of this estimate should be smaller than the one before. The design, execution, and evaluation of N-TPR experiments are repeated until the desired quality of the parameter covariance matrix is obtained.

Optimal control problem for model discrimination

Objective function and optimal control problem To design N-TPR experiments for the purpose of model discrimination, we propose an objective function based on the T-optimality criterion.[19] This criterion assumes that two models, M_1 and M_2, exist and the aim is to design an experiment that maximizes the ability to discriminate between both of them. In our implementation, we assume that parameter estimates are available for both models and we propose the following objective:

$$\max_{u(\tau)} \int_{\tau=0}^{\tau_e} \left(y^{(1)}(u(\tau)) - y^{(2)}(u(\tau)) \right)^2 d\tau. \tag{27}$$

Thus, the optimal N-TPR experiment should give the maximum difference between the output of both rivaling models, $y^{(1)}$ and $y^{(2)}$, and thereby allow to discriminate between both. The resulting optimization problem is given below. Note that it is simpler than the parameter estimation problem in the previous section, especially since sensitivity terms are not needed in the objective function.

$$\max_{u(\tau)} \int_{\tau=0}^{\tau_e} \left(y^{(1)}(\tau) - y^{(2)}(\tau) \right)^2 d\tau$$

Objective function (28)

$$\dot{x}^{(1)} = f^{(1)}\left(x^{(1)}, \vartheta, p^{(1)}\right)$$

ODE constraint, sample comp., model M_1 $\in \Re^{N_s^{(1)}}$ (29)

$$\dot{x}^{(2)} = f^{(2)}\left(x^{(2)}, \vartheta, p^{(2)}\right)$$

ODE constraint, sample comp., model M_2 $\in \Re^{N_s^{(2)}}$ (30)

$$\dot{\vartheta} = u(\tau)$$

ODE constraint, temperature $\in \Re^1$ (31)

$$y^{(1)} = g^{(1)}\left(x^{(1)}, \vartheta, p^{(1)}\right)$$

AE constraint, output signal, model M_1 $\in \Re^1$ (32)

$$y^{(2)} = g^{(2)}\left(x^{(2)}, \vartheta, p^{(2)}\right)$$

AE constraint, output signal, model M_2 $\in \Re^1$ (33)

$$x^{(1)}(\tau = 0) = x_0^{(1)}$$

Initial condition, Eq. 29 $\in \Re^{N_s^{(1)}}$ (34)

$$x^{(2)}(\tau=0)=x_0^{(2)} \quad \text{Initial condition, Eq. 30} \quad \in \Re^{N_s^{(2)}} \tag{35}$$

$$\vartheta(\tau=0)=\vartheta_0 \quad \text{Initial condition, Eq. 31} \quad \in \Re^{1} \tag{36}$$

$$u_{\min} \le u \le u_{\max} \quad \text{Bounds on control variable} \quad \in \Re^{1} \tag{37}$$

$$\vartheta_{\min} \le \vartheta \le \vartheta_{\max} \quad \text{Bounds on temperature} \quad \in \Re^{1} \tag{38}$$

$$x^{(1)}(\tau_e) \le x_{e,\max}$$

$$\text{Complete conversion at end time, model } M_1 \quad \in \Re^{N_s^{(1)}} \tag{39}$$

$$x^{(2)}(\tau_e) \le x_{e,\max}$$

$$\text{Complete conversion at end time, model } M_2 \quad \in \Re^{N_s^{(2)}} \tag{40}$$

This optimization can be used to discriminate between rivaling models that differ with respect to the assumed reaction mechanisms and the number of reactions or the sequence of reactions (parallel or sequential reaction schemes) in systems with single or multiple reactions. In this study, we focus on the discrimination of reaction kinetic laws in systems with a single reaction.

Numerical treatment As with the optimization problem in Section "Optimal control problem for improved parameter estimation," we discretize the DAEs (Eqs. 29–33) according to the orthogonal collocation method on finite elements. Three collocation points per finite element are applied, and the number of elements range between 100 and 300. The discretized problem is implemented in AMPL, and IPOPT and CONOPT are used to solve it. Numerous cases have been solved for competing kinetic models with variations of the parameters of the model M_1.

The optimization procedure is as follows:

- The output signal of the first model, M_1, with given parameters is calculated for a linear TPR experiment with maximum temperature gradient.
- The parameters of the second model, M_2, are estimated by minimizing the squared error between the outputs of both models at the maximum temperature gradient. In this minimization, upper and lower bounds for the parameter values are imposed. The parameters of both models are kept constant throughout the following optimization of the N-TPR experiment.
- Next, a linear TPR experiment is designed by solving Eqs. 28–40 in discretized form with a constant temperature gradient. This yields the best possible linear TPR experiment.
- As a final step, the assumption of a constant temperature gradient is relaxed and an N-TPR experiment is designed using Eqs. 28–40 in discretized form.

Application strategy These N-TPR experiments are used to discriminate between two models with known or estimated parameters. If no parameter estimates exist, then they may be obtained from a priori TPR experiments, which may follow linear temperature profiles or nonlinear TPRs determined in Section "Optimal control problem for improved parameter estimation." Once the parameter values are available, an N-TPR experiment can be designed using Eqs. 28–40 to discriminate between two models. After executing the designed experiment, one should be able to rule out one of the two models. If this N-TPR experiment is not sufficient to statistically disqualify one of both models, then discrimination between these models is not possible with the TPR method alone.

Optimal N-TPR Experiments

The solutions of both optimal control problems are dependent on the choice of model parameters. Therefore, in this study optimizations were carried out over a wide range of parameter combinations. For the sake of brevity, only a few representative solutions are discussed in this section.

Optimal N-TPR experiments for improved parameter estimation

Systems with one reaction—analytical solution Generally, the model equations in Section "The TPR model" cannot be solved analytically. However, for systems with a single reaction, one can manipulate the optimal control problem stated in Section "Optimal control problem" such that some information on the optimal control profile can be obtained. If we assume a system with a single reaction following power law kinetics, then applying Eqs. 15–23 and introducing the variable

$$T=\exp\left(\mathrm{Arr}\cdot\left(\frac{1}{\vartheta^{\mathrm{ref}}}-\frac{1}{\vartheta}\right)\right) \tag{41}$$

gives the following formulation:

$$\max_{u(\tau)} \det \int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau; \qquad p=\{\mathrm{Da},\mathrm{Arr},n,\Theta\} \tag{42}$$

$$\dot{x}=-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot x^n \tag{43}$$

$$\dot{\vartheta}=u \tag{44}$$

$$y=\mathrm{Da}\cdot T\cdot x^n \tag{45}$$

$$\dot{J}_{\mathrm{Da}}=-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot n\cdot x^{n-1}\cdot J_{\mathrm{Da}}-\frac{1}{\Theta}\cdot T\cdot x^n \tag{46}$$

$$\dot{J}_{\mathrm{Arr}}=-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot n\cdot x^{n-1}\cdot J_{\mathrm{Arr}}-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot\frac{\log T}{\mathrm{Arr}}\cdot x^n \tag{47}$$

$$\dot{J}_n=-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot n\cdot x^{n-1}\cdot J_n-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot x^n\cdot\log x \tag{48}$$

$$\dot{J}_\Theta=-\frac{\mathrm{Da}}{\Theta}\cdot T\cdot n\cdot x^{n-1}\cdot J_\Theta+\frac{\mathrm{Da}}{\Theta^2}\cdot T\cdot x^n \tag{49}$$

$$y_{\mathrm{Da}}=\left[\mathrm{Da}\cdot n\cdot x^{n-1}\cdot J_{\mathrm{Da}}+x^n\right]\cdot T\cdot\mathrm{Da} \tag{50}$$

$$y_{\mathrm{Arr}}=\left[\mathrm{Da}\cdot n\cdot x^{n-1}\cdot J_{\mathrm{Arr}}+\mathrm{Da}\cdot\frac{\ln T}{\mathrm{Arr}}\cdot x^n\right]\cdot T\cdot\mathrm{Arr} \tag{51}$$

$$y_n=\left[\mathrm{Da}\cdot n\cdot x^{n-1}\cdot J_n+\mathrm{Da}\cdot x^n\cdot\log x\right]\cdot T\cdot n \tag{52}$$

$$y_\Theta=\left[\mathrm{Da}\cdot n\cdot x^{n-1}\cdot J_\Theta+0\right]\cdot T\cdot\Theta \tag{53}$$

The term Arrhenius, T, occurs linearly in most of the ODEs, so we introduce a new time coordinate, t:

$$\frac{dt}{d\tau} = T \tag{54}$$

This transformed time may be interpreted as a decelerated time, because the reaction and the evolution of all sensitivity functions are accelerated by the Arrhenius term, in the actual time domain. Derivatives with respect to the decelerated time coordinate are noted as x'. The optimization problem in the decelerated time coordinate reads:

$$\max_{u(t)} \det \int_{t=0}^{t_e} \left(\frac{y_p}{T}\right) \cdot \left(\frac{y_p}{T}\right)^T \cdot T \, dt \tag{55}$$

$$x' = -\frac{\mathrm{Da}}{\Theta} \cdot x^n \tag{56}$$

$$\vartheta' = \frac{u}{T} \tag{57}$$

$$J'_{\mathrm{Da}} = -\frac{\mathrm{Da}}{\Theta} \cdot n \cdot x^{n-1} \cdot J_{\mathrm{Da}} - \frac{1}{\Theta} \cdot x^n \tag{58}$$

$$J'_{\mathrm{Arr}} = -\frac{\mathrm{Da}}{\Theta} \cdot n \cdot x^{n-1} \cdot J_{\mathrm{Arr}} - \frac{\mathrm{Da}}{\Theta} \cdot \frac{\log T}{\mathrm{Arr}} \cdot x^n \tag{59}$$

$$J'_n = -\frac{\mathrm{Da}}{\Theta} \cdot n \cdot x^{n-1} \cdot J_n - \frac{\mathrm{Da}}{\Theta} \cdot x^n \cdot \log x \tag{60}$$

$$J'_\Theta = -\frac{\mathrm{Da}}{\Theta} \cdot n \cdot x^{n-1} \cdot J_\Theta + \frac{\mathrm{Da}}{\Theta^2} \cdot x^n \tag{61}$$

$$\frac{y_{\mathrm{Da}}}{T} = \left[\mathrm{Da} \cdot n \cdot x^{n-1} \cdot J_{\mathrm{Da}} + x^n\right] \cdot \mathrm{Da} \tag{62}$$

$$\frac{y_{\mathrm{Arr}}}{T} = \left[\mathrm{Da} \cdot n \cdot x^{n-1} \cdot J_{\mathrm{Arr}} + \mathrm{Da} \cdot \frac{\ln T}{\mathrm{Arr}} \cdot x^n\right] \cdot \mathrm{Arr} \tag{63}$$

$$\frac{y_n}{T} = \left[\mathrm{Da} \cdot n \cdot x^{n-1} \cdot J_n + \mathrm{Da} \cdot x^n \cdot \log x\right] \cdot n \tag{64}$$

$$\frac{y_\Theta}{T} = \left[\mathrm{Da} \cdot n \cdot x^{n-1} \cdot J_\Theta + 0\right] \cdot \Theta \tag{65}$$

Except for Eqs. 57, 59, and 63, these equations can be solved analytically and all sensitivities except y_{Arr} are independent of $T(t)$, which represents the chosen temperature profile. Thus, as long as the sensitivities with respect to the Arrhenius number are neglected, the first two factors of the integrand in the objective function (Eq. 55) are independent of the control variable, u. Moreover, maximizing the objective function is achieved by setting the factor $T(t)$ to its upper bound, by applying the maximum possible temperature gradient. Thus, when the activation energy is not of interest, the highest possible temperature gradient should be applied in single reaction systems.

This result for power law kinetics also applies to single reaction systems with other reaction kinetics in a similar way, and leads to the same maximum temperature gradient policy.

Systems with one reaction—numerical solutions Here we apply the optimization algorithms from Section "Optimal control problem for improved parameter estimation." For a single reaction system, the relative oxygen capacity is always equal to one, so its sensitivity can be omitted from the objective function. We consider only power law kinetics here. In principle, these examinations can be extended to other reaction kinetics as well, and the results are expected to be similar.

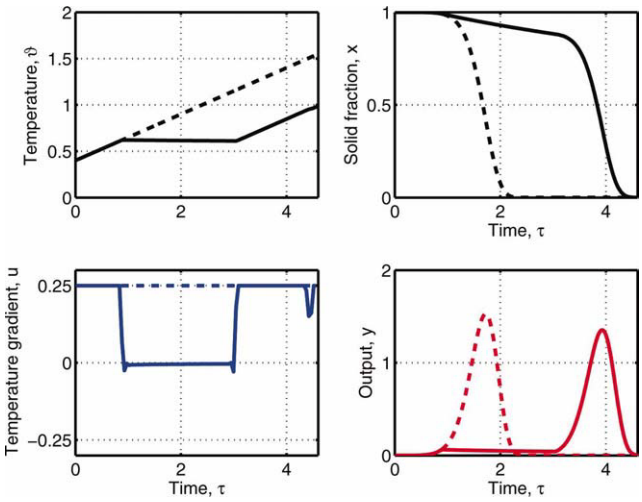

Figure 3. Optimal N-TPR experiment for a single reaction system (power law) with Da = 1.0; Arr = 10; *n* = 1.0; ϑ^{ref} = 0.75.

Dashed lines: profiles from the linear TPR run with maximum temperature gradient. Solid lines: Optimal N-TPR run. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The optimization results are expected to depend on the model parameters selected a priori. Among these, the Damköhler number simply has the effect of shifting the whole TPR signal forward or backward in time. Because this has no qualitative effect on the optimal control profile, we always set this parameter to unity. Also, for single reactions, the oxygen capacity in a single reaction system is also always equal to unity. The two remaining parameters, the Arrhenius number and the order of reaction, are varied systematically in several cases (see Appendix Table A1). Furthermore, several additional cases have been solved to evaluate the impact of parameters such as the number of finite elements, the duration of the N-TPR run and extreme values of the order of reaction. Here, we set the duration of the N-TPR experiment to twice the duration of the linear experiment with maximum temperature gradient.

A typical solution of the optimal control problem is shown in Figure 3. In this example, Arr = 10, which corresponds to an activation energy of 81 kJ/mol, and the order of reaction is equal to unity. The solution is qualitatively similar in most of the other cases. The optimal control profile (lower left diagram) shows that the temperature gradient is at its upper bound, except during a certain time span, where it is close to zero. This isothermal segment is located at the ascending flank of the TPR peak and it delays the occurrence of the peak. Therefore, we refer to this time span of reduced temperature gradient as the "delay phase."

The best possible linear TPR run, which is at maximum temperature gradient, has an objective function value of $F_{\mathrm{lin}} = 9.05 \times 10^{-3}$. The optimal N-TPR experiment has an

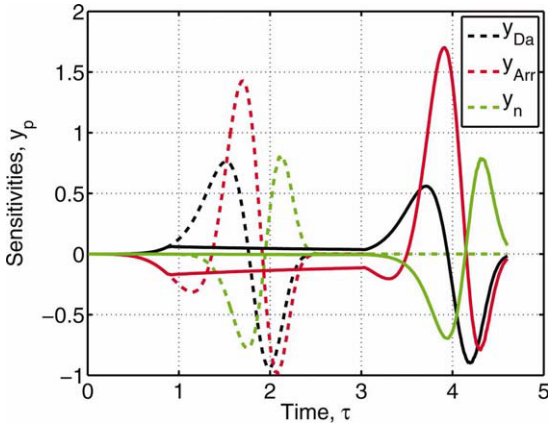

Figure 4. Sensitivity functions for a single reaction system (power law) with Da = 1.0; Arr = 10; n = 1.0; ϑ^{ref} = 0.75. Dashed lines: profiles from the linear TPR run with maximum temperature gradient.

Solid lines: Optimal N-TPR run. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

objective function value of $F^* = 1.15 \times 10^{-2}$. This is an increase by 27%. In other cases, where different values of Arr and n are assumed, the improvement in the objective function from a linear TPR run with maximum temperature to the optimal N-TPR varies between 10 and 70%. Taking into account that the doubling of the TPR run time means an increase of the experiment duration by approximately 25% (including pretreatment, equilibration and cool down phases), this improvement justifies the additional effort.

The improvement of the objective function is due mainly from an increase of the sensitivity with respect to the Arrhenius number, Arr. This corresponds to the expectations that were concluded from the analytical solution in the previous section. The sensitivity functions with respect to the three model parameters (Da, Arr, and n) in the nonlinear and the linear TPR runs are shown in Figure 4. Note that the peak heights of y_n stay almost constant, the peak heights of y_{Da} are slightly lower in the nonlinear run, and the peak heights of y_{Arr} are increased. Although this trend does not seem to be significant, we note that these sensitivity functions are squared and integrated. Thus, a small increase in peak height, especially if the peak is already high, has a significant impact on the objective function value.

In most cases, the typical optimal control profile shows maximum temperature gradients except during a certain period of time at the ascending flank of the TPR peak. For high orders of reaction ($n > 1.5$), a different type of optimal control profile is observed; the delay phase is located at the descending flank of the TPR peak and it is divided into two parts. First, a strongly negative temperature gradient over a short time period is followed by a longer period with almost zero gradient. More detailed investigations have shown that indeed two local optima exist for high orders of reaction: one with a delay phase at the ascending flank and one with a delay phase at the descending flank. For orders of reaction with $n > 1.7$, the profile with the delay phase at the descending flank is better, otherwise the delay phase at the ascending flank (as shown, for example, in Figure 3) leads to a better objective function.

Systems with two reactions To illustrate optimal N-TPR control profiles for systems with multiple reactions, we focus on a system with two consecutive reactions, each with power law kinetics. We assume that both reaction peaks are only about 100°C apart: their reference temperatures are $\vartheta_1^{ref} = 0.75$ and $\vartheta_2^{ref} = 0.85$. Such a system describes any metal oxide that is reduced in two subsequent steps. It has a total of eight parameters that need to be estimated: two Damköhler numbers, two Arrhenius numbers, two orders of reaction, and two oxygen capacities. Usually, one would like to use the determinant of the full sensitivity matrix as the objective function. To simplify the formulation, solution and interpretation of the D-optimal optimum, only the determinant of the sensitivities with respect to both Damköhler numbers and both Arrhenius numbers is considered here. Although this is not the complete determinant, this objective function may still be relevant, as it can be applied to improve the variances and covariances of these four parameters.

$$\max_{u(\tau)} \det \int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau; \qquad p = (\mathrm{Da}_1, \mathrm{Da}_2, \mathrm{Arr}_1, \mathrm{Arr}_2)^T \quad (66)$$

A variety of such cases has been solved with varying Arrhenius numbers and orders of reaction (Appendix Table A2). In addition, the reference temperatures of the reactions were changed in some cases.

As with the single reaction systems, the condition of complete conversion was imposed, Eq. 26, and the end time was set to twice the duration of the corresponding linear run with maximum temperature gradient.

A typical solution of these problems is shown in Figure 5. In this example, the control variable (bottom left diagram) starts at its upper bound, so temperature is increasing (top left diagram) and the first reaction takes place (top right diagram). At about $\tau = 1.7$, when the first reaction's peak is over (see bottom right diagram) and the second reaction rate begins to increase, the control variable is changed to its lower bound for a short period of time, so temperature is decreased. After this, at about $\tau = 2$, the gradient is changed to almost zero for some time. Because the temperature has been decreased previously, the first reaction continues to proceed at low rate during this isothermal phase and reaches almost complete conversion, while the second reaction comes almost to a halt. Towards the end of the TPR run, at about $\tau = 3.9$, the temperature gradient is set back to its maximum value. As a consequence, the TPR peak occurs for the second reaction. The objective function value increases from $F_{lin} = 5.97 \times 10^{-5}$ for the linear TPR run with maximum gradient to $F_{opt} = 7.15 \times 10^{-5}$ for the N-TPR run.

Moreover, the effect of this two-staged delay phase leads to a separation of the signal peaks for the two reactions along the time coordinate. This is also reflected by the elements of the integrated sensitivity product. In the linear experiment with maximum gradient, the determinant is 5.97×10^{-5} and the corresponding matrix is given by:

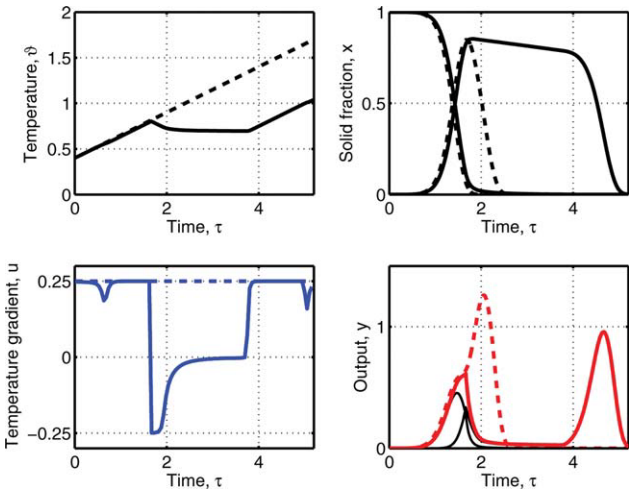

Figure 5. Optimal N-TPR experiment for a system with two reactions (both power law kinetics) with Da = [1.0, 1.0]; Arr = [10, 13]; *n* = [1.0, 1.0]; Θ = [0.75, 0.25]; ϑ^{ref} = [0.75, 0.85].

Dashed lines: profiles from the linear TPR run with maximum temperature gradient. Solid lines: Optimal N-TPR run. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau = \int_{\tau=0}^{\tau_e} \begin{pmatrix} y_{\mathrm{Da1}} \\ y_{\mathrm{Da2}} \\ y_{\mathrm{Arr1}} \\ y_{\mathrm{Arr2}} \end{pmatrix} \cdot \begin{pmatrix} y_{\mathrm{Da1}} \\ y_{\mathrm{Da2}} \\ y_{\mathrm{Arr1}} \\ y_{\mathrm{Arr2}} \end{pmatrix}^T d\tau = \begin{pmatrix} 0.027 & -0.022 & -0.012 & -0.014 \\ -0.022 & 0.253 & 0.015 & 0.130 \\ -0.012 & 0.015 & 0.043 & -0.018 \\ -0.014 & 0.130 & -0.018 & 0.329 \end{pmatrix} \quad (67)$$

For the optimal N-TPR run, the determinant is 7.15×10^{-5} and the corresponding matrix is given by:

$$\int_{\tau=0}^{\tau_e} y_p \cdot y_p^T d\tau = \begin{pmatrix} 0.024 & -0.007 & -0.011 & -0.008 \\ -0.007 & 0.156 & 0.013 & 0.006 \\ -0.011 & 0.013 & 0.043 & -0.011 \\ -0.008 & 0.006 & -0.011 & 0.520 \end{pmatrix} \quad (68)$$

While some of the main-diagonal elements have decreased, instead of increasing, one can determine that the improvements in the objective function are mainly obtained due to the decreases in off-diagonal elements, including $y_{\mathrm{Da2}} \cdot y_{\mathrm{Arr2}}$ and combined elements of parameters from both reactions, $y_{\mathrm{Da1}} \cdot y_{\mathrm{Da2}}$, and $y_{\mathrm{Da1}} \cdot y_{\mathrm{Arr2}}$.

However, such a peak separation was not obtained in all cases (see Appendix Table A2). For certain combinations of n_1 and n_2, a linear TPR run with maximum temperature gradient is optimal, while a secondary delay phase can be observed, for example, in the case shown in Figure 5. Here, the temperature gradient is decreased for a very short period of time at the ascending flank of the first reaction's peak. That indicates that the delay phase, which was optimal in the single reaction systems is still present here, but it is not dominant.

Optimal N-TPR experiments for model discrimination

We now consider the optimization strategy in Section "Optimal control problem for model discrimination" to design N-TPR experiments to discriminate between rivaling single reaction models with different kinetics. With three different reaction kinetics (power law, Avramy-Erofeev and diffusion limitation from Section "The TPR model"), six (ordered) pairs of rival models are possible. The optimization problems for these different model combinations are solved for varying Arrhenius numbers and reaction orders of the first model, M_1 (see Appendix Table A3). All other parameters are set constant: $\mathrm{Da}^{(1)} = 1$, $\Theta^{(1)} = 1$, and $\vartheta^{\mathrm{ref}(1)} = 0.75$.

The optimization results show two different types of optimal control profiles. In some cases, a short delay phase with a strongly negative gradient followed by a longer phase with constant gradient is observed. These control profiles are similar in shape to the peak separating delay phase in Figure 5. In many other cases, the optimal control profile simply has a delay phase with almost zero temperature gradient. Figure 6 shows a typical example of this type of control profile.

The upper right diagram shows the simulated outputs of both models from two linear TPR runs. The output signal of the model M_1 at its maximum temperature gradient is a clean sharp peak ending at about $\tau = 2.2$. Because the parameters of the model M_2 are fitted to this signal, the output profiles of both models are practically identical and the difference between both curves, given by the objective (Eq. 28), is only $F_{\mathrm{lin,ini}} = 3.02 \times 10^{-5}$.

The best possible linear TPR run is obtained at $u = 0.107$, where the difference between both output profiles increases to $F_{\mathrm{lin,opt}} = 0.165$. The two output signals are also shown as the two broader peaks in the upper right diagram. This is a constrained optimum, with an active inequality constraint for complete conversion (Eq. 39); further increasing the experiment time would lead to lower temperature gradients and additional improvement of the objective function value. Although a lower temperature gradient increases the difference between the two profiles by several orders of magnitude, both output signals are still very similar in shape. Conducting such a linear TPR run may produce a measured signal somewhere in between these calculated profiles. Such an experimental result would not allow us to discriminate between the two models.

The output signals of both models from the optimal N-TPR run are shown in the bottom right diagram, with the control profile represented by the solid line in the bottom left. Note that the control profile is at maximum temperature gradient, except during a long delay phase during the middle of the experiment. This leads to an objective function value of $F^* = 0.552$. The N-TPR run not only increases the objective function value by an additional factor of about 3, but also produces two profiles with qualitatively different shapes. Conducting such an experiment should provide a measured signal that allows us to discard one of the two models.

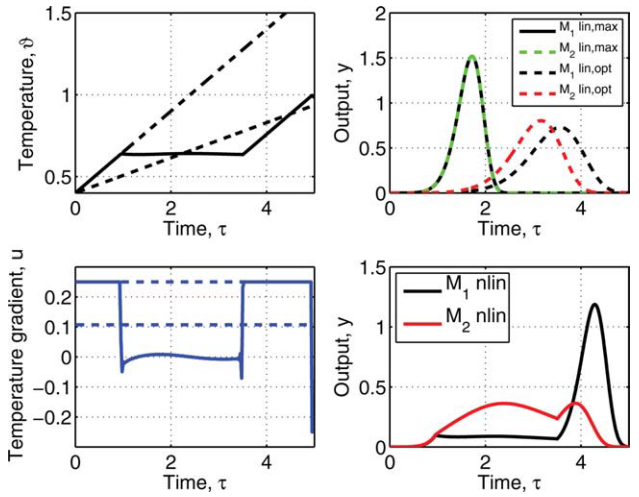

Figure 6. Optimal N-TPR experiment for model discrimination (power law vs. Avramy-Erofeev kinetics).

Upper right corner: Model outputs from linear TPR runs with maximum and optimal temperature gradient; Lower right corner: Model outputs from optimal N-TPR run. The optimal control profile is the blue solid line in the lower left diagram. $Da^{(1,2)} = [1.0, 1.12]$; $Arr^{(1,2)} = [10, 4.27]$; $n^{(1,2)} = [1.0, 0.5]$; $\Theta^{(1,2)} = [1, 1]$; $\vartheta^{ref(1,2)} = [0.75, 0.75]$.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Reduced Problem Formulation

Because of the linear dependence of $u(\tau)$ in Eqs. 17 and 31, both optimization formulations are singular control problems. These ill-conditioned problems are characterized by shallow response surfaces. If the optimal temperature gradient is not at its bounds, then $u(\tau)$ consists of singular arcs. Unique solutions of these profiles are difficult to determine within numerical precision.

Consequently, solving the optimal control problem for an N-TPR experiment is not an easy task, even with advanced methods and an experienced user. The procedure that is necessary to converge towards a useful solution is difficult and not identical for all cases. Initially guessed profiles that work well for some cases may fail for others.

With these characteristics this approach is difficult to implement and automate in a laboratory environment. Nevertheless, from the nature of singular control problems and the solutions in the previous section, we adopt two guidelines for the development of reduced problem formulations.

- Singular control problems can be regularized through the addition of quadratic penalty terms or coarse discretizations of the control profile. Either modification improves the likelihood for better conditioned problems, unique solutions, and faster convergence.
- The solution profiles $u(\tau)$ in Section "Optimal N-TPR Experiments" can be captured reasonably accurately by piecewise constant elements. Choosing such a discretization also leads to better conditioned problem formulations with few degrees of freedom.

These guidelines are applied for the following problem cases.

Reduced problems for parameter estimation

Systems with one reaction The optimal control function for systems with one single reaction typically applies a maximum temperature gradient except during a delay phase, where the temperature gradient is virtually constant but not at its upper bound (see Section "Systems with one reaction—numerical solutions"). These profiles can be approximated by a piecewise constant function, as shown in Figure 7, with three variables: times at the beginning and at the end of the delay phase, τ_1 and τ_2, and the value of the control variable during the delay phase, u_1.

This reduced optimization problem has far fewer degrees of freedom and may be solved with standard software such as MATLAB. In this environment, the DAE constraints are integrated numerically at each function evaluation. In our implementation, we approximate the gradients by finite differences and use the MATLAB function "fmincon," which uses an SQP algorithm with a BFGS update of the Hessian. It allows implementing the upper and lower bounds of the optimization variables as inequality constraints.

The integration of the ODE constraints is conducted using *ode15s*, an implicit Runge-Kutta algorithm of variable order. It uses an adapted step width in time, and the resulting discretization in time is a bit finer than in the solution of the full problems. Due to these different discretization schemes, the objective function values in the full and in the reduced problem may differ, but optimal profiles can be compared directly.

The reduced problems are typically solved within 30–50 optimization steps, which usually takes a few CPU minutes (Intel Core2 Duo CPU E6850, 3.00 GHz). Convergence is reliable in more than 90% of the cases; only the choice of the initial point requires some input by the user.

The reduced problems have been solved for the same combinations of the Arrhenius number and the order of reaction as for the full problem. In most of the cases, the optimal control profiles obtained from the reduced problem is very similar to the results from the corresponding full problem, and the nonlinear experiments show similar improvements compared to their linear counterparts. In particular, the delay phase is always located at the beginning of the ascending flank of the TPR signal.

For illustration, we reconsider the case from Section "Systems with one reaction—numerical solutions." Here, the solid lines in Figure 8 show a typical result for the reduced problem formulation. For comparison, the solution from the corresponding full problem is shown in dashed lines and the two solutions have outputs and temperature profiles that are almost identical. This observation applies to the majority of the solved cases, so this problem reduction seems to be appropriate. For this case the objective function for the full problem increases by 27%, from $F_{lin} = 9.05 \times 10^{-3}$ to $F^* = 1.15 \times 10^{-2}$. For the reduced problem, it increases by 45%, from $F_{lin} = 1.39 \times 10^{-2}$ and $F^* = 2.02 \times 10^{-2}$.

Systems with two reactions The shape of the delay phase that separates the two peaks is different from the typical shape of the delay phase found with single reaction systems

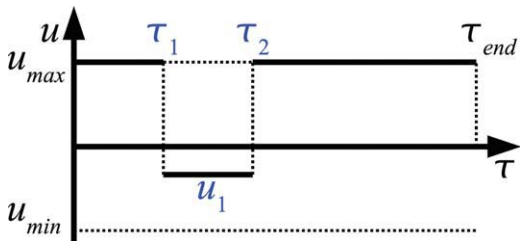

Figure 7. Schematic control profile for the reduced problem for single reaction systems.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

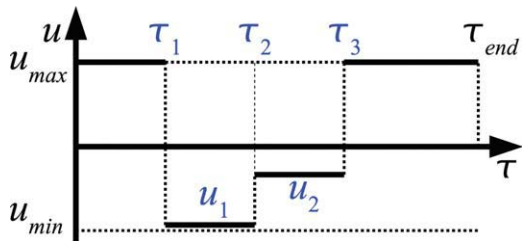

Figure 9. Schematic control profile for the reduced problem for systems with two reactions.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

(Section "Systems with one reaction—numerical solutions"). The profile in Figure 5 includes a strongly negative temperature gradient followed by a temperature gradient close to zero. Therefore, the delay phase is split into two parts, introducing one new segment to the reduced problem. This profile is depicted in Figure 9 with segments u_1 and u_2, respectively.

As shown in Appendix Table A2, many reduced problem cases have been solved, using the same parameter combinations that were applied in the solutions of the full problem. A typical solution for a system with Arr = [10, 10], n = [1.0, 1.0] is shown in Figure 10 (right). It has a delay phase with a cooling period located between the two peaks. The control profile and the output signal closely approximate the typical peak separating solution of the corresponding full problem, which is shown in dotted lines in the diagram. Similar solutions are observed in many cases with different parameters.

However, depending on the initially chosen control profile, the optimization converges to a second local optimum (Figure 10, left). This optimal control profile has a delay phase at the ascending flank of the first peak. No such delay phase was observed in the solutions of the full problems, so this second local optimum occurs due to the reduction of the control profile. With regard to the objective function, the control profile without peak separation (Figure 10 left) is better than the solution with peak separation (right diagram). This applies to many other cases in Appendix Table A2.

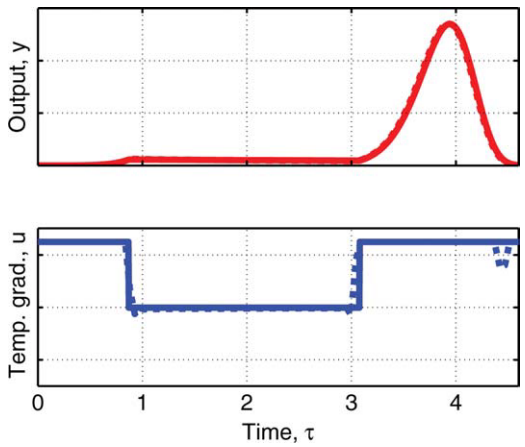

Figure 8. Optimal N-TPR experiments for a single reaction system (power law) with Da = 1.0; Arr = 10; n = 1.0; ϑ^{ref} = 0.75 (c.f. Section "Systems with one reaction—numerical solutions").

Dashed lines: solution of the full problem; solid lines: Solution of the reduced problem. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

With two locally optimal solutions, the question arises, which of both solutions should be applied in an N-TPR experiment. At first glance, it seems to be reasonable to apply the control profile which gives the better objective function value. In the case in Figure 10 that would be the solution shown in the left diagram. There is, however, another practical aspect that needs to be considered: the sensitivity or robustness of the solution. In practice, the system parameters are not exactly known, and the control profile cannot be realized precisely as described in the optimal solution, even if obtained from a reduced problem. If the objective function is very sensitive to changes in the control profile, then it is very likely that small errors in the experimental procedure will result in a very bad experiment. Thus, it is advantageous if the applied solution has a low sensitivity with respect to the control profile. These sensitivities are represented by the eigenvalues of the Hessian matrix of the solution. They are listed in Table 1 for the two solutions shown in Figure 10. Obviously, the solution with a delay phase at the first peak is much more sensitive than the solution with peak separation. This argument advocates the application of the peak separating solution in the right diagram, because it is more robust, although it may not be globally optimal in all cases. Because in such uncertain systems robustness may be an important issue, it should be considered in the optimization, for example, through a robustness term in the objective function. This aspect needs further detailed consideration and is left as a subject for future studies.

Another approach to solve the problem of multiple optima is to apply more complex control profiles. In the case here, a control profile with two distinct delay phases could be chosen. The first delay phase is supposed to be located at the ascending flank of the first peak, and the second delay phase, which should have a cooling phase and an approximately isothermal phase, should separate the two peaks. This control profile combines the two optimal profiles in Figure 10 and should lead to unique optima. Further investigation will be considered in future studies.

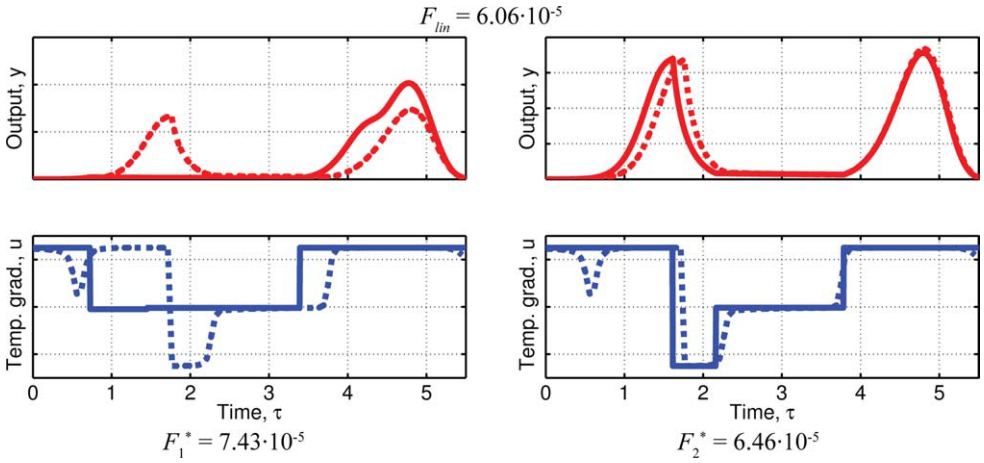

Figure 10. Example of two local optima of the reduced problem.

The dotted lines indicate the solution of the full problem and the solid lines are the solutions of the reduced problems. Da = [1.0, 1.0], Arr = [10, 10], n = [1.0, 1.0], Θ = [0.75, 0.25], ϑ^{ref} = [0.75, 0.85]. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Reduced problems for model discrimination

The solutions of the full problems for model discrimination in Section "Optimal N-TPR experiments for model discrimination," typically show a control profile with a single delay phase where the temperature gradient is close to zero. In some cases, an additional delay phase is also included with a strongly negative temperature gradient. In the reduced problem, we ignore the possible occurrence of this cooling phase and instead favor a simple profile with only one delay phase, as shown in Figure 7.

A typical solution of the reduced problem is shown in Figure 11 together with the solution of the corresponding full problem. The objective function value is increased from $F_{\text{lin,ini}} = 3.03 \times 10^{-5}$ for a linear TPR run with maximum temperature gradient to $F^* = 0.543$ at the solution of the reduced problem. The function values are usually very similar to those obtained from the full problems. This is aided by the absence of sensitivity terms in the objective function; these may have very strong curvature and are therefore sensitive to discretization errors.

The results for other combinations of models show similar trends and are not shown here. Also in those cases, where the full solution shows a double delay phase, the output signals and the objective function values of the full and the corresponding reduced problem are very similar. This means that the proposed problem reduction is appropriate. Even in some cases, where no solution could be found for the full problem (see Appendix Table A3), the reduced problem was solved successfully. However, as seen in Appendix Tables A1–A3 some cases remain where no nonlinear solution could be found. This suggests that in these cases, no improvement of the objective function value is possible and a linear TPR experiment is optimal.

The initialization of the optimization problem is simple and reliable in most cases. A good initial guess is any control profile with a delay phase where the temperature gradient is close to zero and which has a higher objective function value than the linear experiment with maximum gradient. If such an initial solution can be found, the optimization converges within 10–20 optimization steps. These problems are usually solved within a few minutes of CPU time.

Conclusions

Optimal control problems for the design of nonlinear TPR experiments have been set up and solved for a wide range of parameters and for several different kinds of systems. Selected examples of their solutions are presented in Section "Optimal N-TPR Experiments." They show that N-TPR can improve the quality of parameter estimates obtained from TPR experiments. In some cases, the improvements are significant, while in other, they are only small and a linear TPR experiment is almost as good. For systems with more than one reaction, optimal control profiles often yield so-called peak separation, where the measurement signals of both reactions are separated in time. This reduces the covariances of the parameter estimates. The design of optimal N-TPR experiments for model discrimination yields strong improvements in model discrimination capabilities.

The full problem formulation is a singular control problem that may be difficult to solve. Nevertheless, the optimal control profiles can often be approximated by piecewise constant functions. In accordance with the analytical solution in Section "Systems with one reaction—analytical solution," the control variable is usually at its upper bound, except during

Table 1. Eigenvalues of the Hessian Matrices of the Solutions Shown in Figure 10

Position of delay phase	Ascending flank of first peak	Between both peaks
Eigenvalues of Hessian matrix	$\lambda = \begin{pmatrix} 5.2 \cdot 10^{3} \\ 3.7 \cdot 10^{5} \\ 2.5 \cdot 10^{6} \\ 9.5 \cdot 10^{6} \\ 1.7 \cdot 10^{14} \end{pmatrix}$	$\lambda = \begin{pmatrix} 1.3 \cdot 10^{-3} \\ 6.5 \cdot 10^{1} \\ 1.5 \cdot 10^{2} \\ 9.0 \cdot 10^{2} \\ 1.7 \cdot 10^{4} \end{pmatrix}$

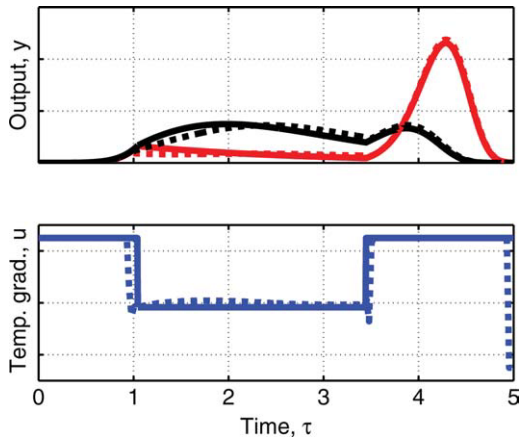

Figure 11. Solution of the reduced problem for model discrimination (power law vs. Avramy-Erofeev kinetics).

Dotted lines: Solution of the corresponding full problem (see Figure 6); solid lines: Solution of the reduced optimization problem. Black lines: Output model M_1; Red lines: output model M_2; Blue lines: control profiles. Da = [1.0, 1.12]; Arr = [10, 4.27]; n = [1.0, 0.5]; Θ = [1, 1]; ϑ^{ref} = [0.75, 0.75]. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

so-called delay phases, where it is close to zero or near its lower bound. Consequently, the reduction of the optimal control problems to optimization problems of lower dimension is possible and successful in most cases. The solutions of the reduced problems usually approximate the solutions of the corresponding full problems very well. In addition, for many parameter combinations that could not be solved in the full formulation, a good reduced solution could be obtained.

In one group of cases, the reduced problems had at least two local optima. This was not observed with the solutions of the corresponding full problems. This point may be amended by applying different, more complex control profile schemes with more than two delay phases.

Numerical convergence of the optimization algorithm for the reduced problems is quite reliable. The only critical point is the identification of a suitable initial guess for the profiles. In some cases, this may be difficult and require some additional input by the user. Besides this, the procedure can be made to work in a fully automated way.

The solutions of the reduced problem are control profiles with piecewise constant temperature gradients. They resemble a sequence of linear TPR experiments, which, in principle, can be applied in TPR devices. Nevertheless, depending on the device, changes to the control software may be necessary so that sequences of linear temperature profiles can be applied. Moreover, to realize negative temperature gradients, as needed for multireaction systems (see Figure 10 right), an appropriate cooling device is required.

Finally, imperfections such as inhomogeneous spatial temperature distribution or nonideal temperature control lead to deviations from the desired, optimal temperature profile. This affects both linear and nonlinear TPR experiments alike, so the N-TPR experiments are still expected to be superior. A deeper analysis of this issue, which concerns the robustness of the optimal nonlinear control profiles, is left as subject for further studies.

Acknowledgment

A part of this work was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD).

Notation

The formulation of the optimization problems and the presentation of their results have been done in terms of dimensionless parameters.

Arr_j = Arrhenius number of reaction j
B = sensitivity matrix of measurements
Da_j = Damköhler number of reaction j
E_j = activation energy of reaction j, [J/mol]
f = right-hand side of state equation (Eq. 1)
g = right-hand side of output equation (Eq. 5)
J = Jacobian, parameter sensitivity of states
n_j = order of reaction j
N_s = number of species
N_p = number of model parameters
p = vector of model parameters to be estimated
R = universal gas constant, [$J/mol/K$]
R_j = rate of reaction j
r_j = concentration dependent term of R_j
T = Arrhenius term
T^0 = reference temperature, [K]
t = decelerated time
u = control variable, temperature gradient
V_0 = a priori covariance of parameter estimates
x_i = fraction of solid species i
y = output variable, measurement signal
y_p = parameter sensitivity of output variable
Θ_j = relative oxygen capacity of reaction j
ϑ = temperature
λ = Eigenvalues of Hessian matrix
$\nu_{i,j}$ = stoichiometric coefficient of species i in reaction j
Π = covariance matrix of measurements
τ = Time

Subscripts

i = solid species
ini = initial profile
j = reaction
lin = linear TPR experiment
min = lower bound
max = upper lound
p = derivative w.r.t. model parameters
x = state
0 = initial value at $\tau = 0$

Superscript

ref = reference
(1),(2) = model 1, 2
θ = standard
$*$ = optimum

Literature Cited

1. Kissinger HE. Reaction Kinetics in Differential Thermal Analysis. *Anal Chem.* 1957;29:1702–1706.
2. Hurst NW, Gentry SJ, Jones A. Temperature Programmed Reduction. *Catal Rev—Sci Eng.* 1982;24:233–309.
3. Heidebrecht P, Galvita V, Sundmacher K. An alternative Method for Parameter Identification from Temperature Programmed Reduction (TPR) data. *Chem Eng Sci.* 2008;63:4776–4788.
4. Wimmers OJ, Arnoldy P, Moulijn JA. Determination of the Reduction Mechanism by Temperature-Programmed Reduction: Application to small Fe_2O_3 Particles. *J Phys Chem.* 1986;90:1331–1337.
5. Piotrowski K, Mondal K, Lorethova H, Stonawski L, Szymański T, Wiltowski T. Effect of Gas Composition on the Kinetics of Iron Oxide Reduction in a Hydrogen Production Process. *Int J Hydrogen Energy.* 2005;30:1543–1554.

6. Pineau A, Kanari N, Gaballah I. Kinetics of Reduction of Iron Oxides by H_2 – Part I: Low Temperature Reduction of Hematite. *Thermochim Acta*. 2006;447,89–100.
7. Vyazovkin S, Dollimore D. Linear and Nonlinear Procedures in Isoconversional Computations of the Activation Energy of Nonisothermal Reactions in Solids. *J Chem Inf Sci*. 1996;36:42–45.
8. Wimmers OJ. Analytical Expressions for Temperature Programmed Patterns using Solid-State Kinetics. *Thermochim Acta*. 1985;95:67–72.
9. Levenspiel O. *The Chemical Reactor Omnibook*. Corvallis: OSU Book Stores, 1996.
10. Vyazovkin S, Wight CA. Kinetics in Solids. *Annu Rev Phys Chem*. 1997;48:125–149.
11. Avrami M. Kinetics of phase change I – General theory. *J Chem Phys*. 1939; 7:1103–1112.
12. Kanervo JM, Krause AO. Kinetic Analysis of Temperature-Programmed Reduction: Behavior of a CrO_x/Al_2O_3 Catalyst. *J Phys Chem B*. 2001;105:9778–9784.
13. Pukelsheim F. *Optimal design of experiments*. New York: Wiley, 1993.
14. Atkinson AC, Donev AN, Tobias RD. *Optimum experimental design*. Oxford: Oxford University Press, 2007.
15. Bard Y. *Nonlinear parameter estimation*. London: Academic Press, 1974.
16. Biegler LT. *Nonlinear programming: concepts, algorithms and applications to Chemical Processes*. Philadelphia: SIAM, 2010.
17. Fourer R, Gay DM, Kernighan BW. *AMPL: a modeling language for mathematical programming*, 2nd ed. Pacific Grove: Brooks/Cole – Thomson, 2003.
18. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program*. 2006;106:25–57.
19. Atkinson AC, Fedorov VV. Optimal design: experiments for discrimination between several Models. *Biometrika*. 1975;62:289–303.

Appendix

Table A1. Solved Cases for a System with One Single Reaction

n	0.7	1.0	1.5
Arr			
5	+	+	+
10	+	+	+
15	+	+	+

"+": cases where the optimization converged to an optimum.

Table A2. Solved Cases for a System with Two Reactions

		$n_1 \vert n_2$				
		1.0|1.0	1.0|1.5	1.5|1.0	1.0|0.7	0.7|1.0
$Arr_1 \vert Arr_2$	10|10	+	−	+	−	−
	10|13	+	−	+	−	+
	13|10	−	−	+	−	−
$\vartheta_2^{ref} = 0.95$						
$Arr_1 \vert Arr_2$	10|13	+	+	+	−	+

"+": cases where the optimization converged to an optimum with peak separation; "−": cases where no convergence was obtained and the linear TPR experiment is optimal or nearly optimal.

Table A3. Solved Cases for Model Discrimination

M_1: Power law M_2: Avramy-Erofeev			
n_1	0.5	1.0	1.5
Arr_1			
5	+	+	+
10	+	+	+
15	+	+	+
M_1: Avramy-Erofeev M_2: Power law			
n_1	1/2	2/3	3/4
Arr_1			
5	+	+	+
10	+	+	+
15	lin	lin	lin
M_1: Diffusion limited M_2: Power law			
n_1	1.5	2.0	3.0
Arr_1			
5	+	+	+
10	+	+	−
15	+	−	−

"+": cases where the optimization converged to an optimum; "−": cases where no convergence was achieved; "lin": cases where a linear experiment is optimal.

Manuscript received July 2, 2010, and revision received Oct. 19, 2010.