Chemical Engineering Journal 433 (2022) 134121

Contents lists available at ScienceDirect

Chemical Engineering Journal

journal homepage: www.elsevier.com/locate/cej

A model-free sparse approximation approach to robust formal reaction kinetics

Frederic Felsen, Karsten Reuter, Christoph Scheurer *

Fritz-Haber-Institut der Max–Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany
Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstr. 4, D-85747 Garching, Germany

ARTICLE INFO

Keywords:
Kinetic modeling
Data driven kinetic analysis
Pareto optimization

ABSTRACT

Accurate and transferable models of reaction kinetics are of key importance for chemical reactors on both laboratory and industrial scale. Usually, setting up such models requires a detailed mechanistic understanding of the reaction process and its interplay with the reactor setup. We present a data driven approach which analyzes the influence of process parameters on the reaction rate to identify locally approximated effective rate laws without prior knowledge and assumptions. The algorithm we propose determines relevant model terms from a polynomial ansatz employing well established statistical methods. For the optimization of the model parameters special emphasize is put on the robustness of the results by taking not only the quality of the fit but also the distribution of errors into account in a multi-objective optimization. We demonstrate the flexibility of this approach based on artificial kinetic data sets from microkinetic models. This way, we show that the kinetics of both the classical HBr reaction and a prototypical catalytic cycle are automatically reproduced. Further, combining our approach with experimental screening designs we illustrate how to efficiently explore kinetic regimes by using the example of the catalytic oxidation of CO.

1. Introduction

Detailed knowledge about reaction kinetics for large and industrially relevant reaction networks are the basis for a rational design and efficient operation of chemical reactors. However, when studying a new catalytic system the underlying mechanism is in general unknown. Concentration profiles obtained at different reaction conditions contain valuable information about the (inter-) dependence of reactants and products. Extracting these dependencies to gain a deeper understanding of the catalytic mechanism is of general importance in heterogeneous as well as in homogeneous thermal catalysis, on which we will focus here without loss of generality. As illustrated in Fig. 1 the central observable in form of a concentration profile is independent of the catalytic process at hand. In homogeneous catalysis, reactions are usually performed in a batch reactor. Probing the reaction mixture at different points in time and quenching the respective samples directly yields a concentration profile. When studying heterogeneous catalysts in a gas flow reactor the same information is contained in profiles of concentrations versus residence time in the reactor bed. The residence time can be adjusted by either modulating the gas flow through the reactor or by taking samples at different points along the reactor tube.

Researchers have several system parameters at their disposal to influence the outcome of the reaction. These comprise feed concentrations of the reactants, temperature, pressure, as well as a multitude of reactor-specific technical parameters. By recording concentration time profiles under different reaction conditions, information regarding the reaction kinetics can be deduced. In mean-field chemical kinetics the rate of a given reaction is generally expressed in form of rate laws for the time-dependent concentrations $[\mathrm{X_i}](t)$ of the $1 < i < N$ reaction components (reactants and products),

$$\frac{d[\mathrm{X_i}]}{dt} = \sum_j k_j M_{ij} \prod_m \left([\mathrm{X_m}]^{s_{jm}}\right) \quad , \tag{1}$$

where k_j are rate constants, M_{ij} are the elements of the stoichiometric matrix and s_{jm} are the stoichiometric coefficients of the reactants [1,2]. Even for complex reaction mechanisms, these rate laws can often be asymptotically approximated by effective rate equations with only a few dominating terms. An example would be

$$\frac{d[\mathrm{P}]}{dt} \propto [\mathrm{A}]^{\nu_\mathrm{A}} \cdot [\mathrm{B}]^{\nu_\mathrm{B}} \quad , \tag{2}$$

where [P] refers to the concentration of the product depending on the reactant concentrations [A] and [B]. The powers ν_A, ν_B are then referred to as (partial) apparent reaction orders. These simplified equations are generally only valid in certain reaction regimes. A famous example for such effective equations is the Lindemann mechanism [3]. It describes

* Corresponding author at: Fritz-Haber-Institut der Max–Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany.
E-mail address: scheurer@fhi-berlin.mpg.de (C. Scheurer).

https://doi.org/10.1016/j.cej.2021.134121
Received 12 October 2021; Received in revised form 7 December 2021; Accepted 9 December 2021
Available online 1 January 2022

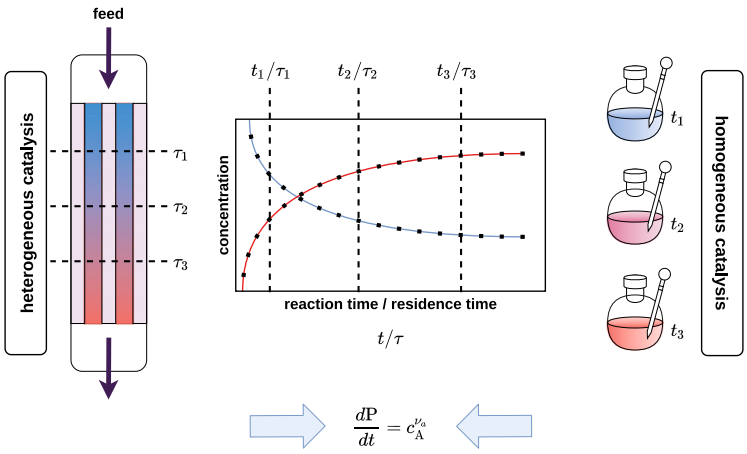

Fig. 1. Concentration–time profiles (e.g. from flow (left) or batch (right) measurements) as central quantity in the kinetic analysis of catalytic reactions.

the self-activation of gaseous species with a subsequent reaction of the activated species. Under low pressure conditions this reaction can be described by an effective second-order law, while under high pressure the reaction is dominated by the subsequent step leading to an effective first-order dependency.

A variety of methods has been developed in order to deduce such effective kinetic equations from concentration and rate versus time data. Particularly in heterogeneous catalysis, such methods often start off from a detailed microkinetic model [4], either in mean field approximation [5] or within surface arrangement-resolved kinetic Monte Carlo [6]. Establishing such a microkinetic model requires a detailed mechanistic understanding of the reaction though, the lack of which is typically one of the prime motivations for establishing effective kinetic mechanisms in the first place. Furthermore, the more complex the reaction network, the more kinetic parameters like activation barriers or exponential prefactors need to be determined on the basis of experimental data. This can lead to questionable inference on the model parameters or even overfitting due to the lack of available data [7]. In homogeneous catalysis, Blackmond [8] established the concepts of reaction progress kinetic analysis (RPKA) and graphical rate equations in 2005. This analysis investigates the dependence of the reaction rate on regime-determining parameters like reactant concentrations. Building on these ideas, methods like the variable time normalization analysis (VTNA) by Burés et al. [9–11] and nonlinear fitting techniques [12,13] have been introduced. They come with the advantage of directly operating on the concentration time profiles to obtain effective reaction orders in the individual components. Additionally, such analyses have proven to be robust against sparse and noisy data, straightforward to interpret and simple to perform [9]. There are, however, also some severe caveats. Notably, a separate analysis of the kinetics in each reactant completely neglects any potential interaction between multiple species. In experimental design [14], parameters which influence the outcome of an experiment and can be controlled by the experimenter are referred to as factors. In order to estimate the interactions of these factors, a proper factorial design is required that varies multiple factor settings at the same time. Only if this or any prior knowledge of the system shows that any potential interaction can be excluded, a separate analysis of all factors is valid.

In general, many of the existing kinetic modeling techniques thus suffer from *a priori* model assumptions. The latter are prone to introduce systematic errors and thus a non-normal error distribution in the model predictions, in contrast to the Gaussian distributed error that would be expected from statistical measurement errors. If the model structure does not fit the physical problem underlying the data, the distribution of prediction errors will correspondingly be skewed. This may be due to an important reaction pathway missing in the assumed mechanism, but also unnecessarily complicated models can be a source of error. An analysis of the residual distribution would readily reveal such issues, but unfortunately it is often neglected. In cases where the residuals are indeed provided [7,13], systematic deficiencies of the models become obvious immediately.

Within this perspective, we here introduce an algorithm for the systematic investigation of formal kinetics from experimental data. Effective rate laws are determined from concentration profiles acquired at different reaction conditions. Performing these measurements according to statistical design [14] allows the algorithm to quantitatively evaluate interactions between different experimental parameters. Centrally, applying well established statistical methods [15–17] we directly incorporate the goal of normally distributed errors into the optimization. At the same time, the algorithm determines relevant terms in the model function automatically [18,19], making the approach to a certain degree free of *a priori* model assumptions.

2. Theory

The apparent reaction order can be understood as a measure of the sensitivity of the reaction rate to a change in reactant concentration [9]. Similar, sensitivity parameters can also be defined for other reaction conditions and are sometimes interpreted as physical quantities. The algorithm, proposed below, is designed to determine effective rate laws by optimizing these sensitivity parameters. This basic idea is closely related to VTNA. In VTNA, the determination of the order in a certain reactant starts with a set of concentration versus time profiles at different reactant concentrations. Then, if the time axes for all profiles are normalized by the integral over the reactant concentrations raised to the correct power, all profiles should coincide. Looking at Eq. (2), we can identify this power (or exponent) as the apparent reaction order. It is the sensitivity parameter of the reaction rate with respect to the reactant concentration. In VTNA, the correct value is typically obtained manually in a trial and error process or in a graphical manner, as coinciding curves can be easily identified by visual inspection.

In contrast, our algorithm uses a more quantitative criterion to identify the correct sensitivity parameters. VTNA works, because the relation between the change in product concentration and the reactant concentration becomes linear, if raised to its respective reaction order. How linear a relation between two variables is, can easily be analyzed by performing linear regression and observing the quality of the resulting model. A well established method, based on the same principle, is the Arrhenius plot. There, both temperature and rates are scaled by a nonlinear transformation revealing a linear relationship with a slope equal to the activation barrier of the process. In the case of the Arrhenius problem, the nonlinear transformations are known from the underlying theory and we are interested in the slope of the resulting linear model. In contrast, if we want to determine effective

rate laws, it is known that there should exist a set of transformations, which leads to a linear relationship. Yet, this set of transformations is not easily accessible without detailed knowledge of the reaction mechanism. Notwithstanding, knowing that a linear relationship should exist, the quality of the regression model can be used as a measure of linearity and we can formulate an optimization problem to automatically identify reaction orders and sensitivity parameters for various factors influencing the reaction. These sensitivity parameters are the key quantity in our approach, as they are supposed to capture all nonlinear effects encountered in reaction kinetics.

We assign one of these sensitivity parameters to every factor, which influences the rate of the reaction. For reactant concentrations, these parameters correspond to their respective partial reaction orders. And, as we will show in a later section, also the apparent barrier can be identified as such a sensitivity parameter by rescaling the reaction temperatures. In the regression model we want to set up, these parameters are the exponents, which introduce nonlinear effects to every factor. So, if we vary the concentration of reactant A ([A]) in a reaction, the corresponding quantity, which enters our model, would be

$$\{\mathrm{A}\} = [\mathrm{A}]^{\alpha_\mathrm{A}} \quad , \tag{3}$$

where $\{\mathrm{A}\}$ denotes the transformed reactant concentration and α_A is the corresponding sensitivity parameter. As this is a type of power transformation [16], we will also refer to these sensitivity parameters (α) as transformation parameters. If reactant A contributes linearly to the reaction rate, e.g. in a simple first order reaction, the corresponding α would be 1. For a second order reaction in A, on the other hand, the optimal α should be close to 2.

2.1. Model function

As long as we are only interested in a single influencing factor, for example we only want to vary the concentration of one reactant, a linear model is sufficient to describe the concentration profile, given the correct scaling of the input and output variables. In a more general case, however, we have multiple reactants and additional factors like pressure or temperature which influence the rate of a chemical reaction. Further, we would like to investigate the influence of these multiple factors on the production rate of possibly even multiple product species. Therefore, we need a more complex model ansatz, which allows for possible interactions and higher order terms. For this reason, our algorithm relies on polynomial functions to describe effective kinetic models. Let us consider a simple reaction with two reactants A and B forming the product P. The only factors we want to account for are the concentrations of these two reactants. In that case, the corresponding polynomial would be

$$\frac{d[\mathrm{P}]}{dt} = \mathrm{const.} + k_1 \cdot \{\mathrm{A}\} + k_2 \cdot \{\mathrm{B}\} + k_3 \cdot \{\mathrm{A}\} \cdot \{\mathrm{B}\} + k_4 \cdot \{\mathrm{A}\}^2 + k_5 \cdot \{\mathrm{B}\}^2 + \cdots \tag{4}$$

with k_i the unknown coefficients of the respective polynomial terms, which need to be determined in the regression. Note that this is now a function of the transformed factors, meaning the transformation parameters (α) are already contained in the terms of this sum. Such a polynomial contains not only higher powers of the input variables but, most importantly, also the necessary interaction terms between multiple factors, meaning e.g. the product $\{\mathrm{A}\} \cdot \{\mathrm{B}\}$. At the same time, polynomial regression is still linear in the unknown parameters, meaning that the quality of the model can still be used as a measure of adequacy for the determined sensitivity parameters. However, being too rigid with the functional form of our model may introduce systematic errors. To account for this we use the least absolute shrinkage sparsification operator (LASSO) [18,19], to let the algorithm itself decide, which polynomial term to include in the final model. LASSO is a modification to the standard linear regression, which, by introducing a regularization term, is able to set regression coefficients to 0. This way, all terms which do not significantly contribute to the solution based on the available data, are automatically canceled out. Further, the remaining (non zero) coefficients not only tell us, which polynomial terms contribute to the effective model. The magnitude of the LASSO coefficients give us a measure of the relative importance of different polynomial terms or, in this context, kinetic processes. Additionally, the sign of the coefficient indicates a positive or negative influence on the reaction rate.

Despite selecting only the significant terms from the polynomial ansatz, we do not want to expand Eq. (4) to arbitrary order, as this would result in a huge amount of available terms to choose from. The final result will depend on the order up to which these ansatz terms are expanded and not every expansion order is reasonable for every problem. Hence, we need to define some guidelines to choose an appropriate ansatz order. First of all, the order of the ansatz terms should generally not exceed the number of factors in our reaction. If we only consider two reactant concentrations for our model, an expansion up to 3rd order would be pointless, as we cannot define any 3-factor interactions in that case. Including e.g. temperature as a third factor, such 3rd order terms could become relevant and should be included in the ansatz. Further, we only consider interactions (or multilinear) terms, meaning product terms of multiple different factors like $\{\mathrm{A}\} \cdot \{\mathrm{B}\}$ as opposed to the purely quadratic ($\{\mathrm{A}\}^2$) or higher power ($\{\mathrm{A}\}^3, \{\mathrm{A}\}^4, \ldots$) terms. The reasoning behind this is that in some cases these terms are already captured by the power transformation in the original factors. For example, a quadratic term with transformation parameter $\alpha = 1$ is equivalent to a linear term with $\alpha = 2$

$$\left([\mathrm{A}]^1\right)^2 = \{\mathrm{A}\}^2_{\alpha_\mathrm{A}=1} = \{\mathrm{A}\}_{\alpha_\mathrm{A}=2} = \left([\mathrm{A}]^2\right)^1 \quad . \tag{5}$$

This, of course, does only hold for very simple systems with only one polynomial term containing [A]. There are exceptions, where including the purely quadratic terms can be beneficial though, for example, when the target property (here the reaction rate) shows an extremal (stationary) point within the region, we want to describe and we will further illustrate and discuss this point below. In general, however, for an initial analysis we neglect these terms. We focus on the remaining pair and higher-order interaction terms. These are especially relevant for factors which are reactant concentrations. In classical collision theory [20], the collision probability of multiple species directly depends on the product of their respective concentrations. This directly transfers into chemical kinetics as a collision (or at least spatial proximity) is an essential requirement for any kind of reaction to take place. This is the reason, why these interaction terms often appear in kinetic rate laws and hence we need to consider them in our model ansatz. The probability of many-body collisions, of course, depends on the type of reaction at hand. Orders of four or higher, however, do usually not contribute significantly, and thus provide a natural upper bound. Another important point is the amount and type of data available, limiting the number of parameters, which can reasonably be estimated. Adding higher order terms also introduces additional degrees of freedom, giving the algorithm more flexibility. This can result in a more accurate model, which in turn will be less interpretable due to a multitude of selected terms. There are, thus, various things to consider when choosing the order of the ansatz polynomial. In practice, a reasonable rule of thumb is to choose a 2nd order ansatz for two input features and a 3rd order ansatz for more than two features. A schematic, showing how the initial factors translate into the target property for the simple example $\mathrm{A} + \mathrm{B} \rightarrow \mathrm{P}$, is given in Fig. 2.

2.2. Cost function

For the optimization of the sensitivity parameters, a quantitative objective is necessary. We want to obtain a statistically robust model, which at the same time achieves a reasonable fit of the provided data. The quality of the fit could be quantified by simply looking at the residual norm of the model. However, at the same time, we want to make sure, that the LASSO regression does not introduce systematic

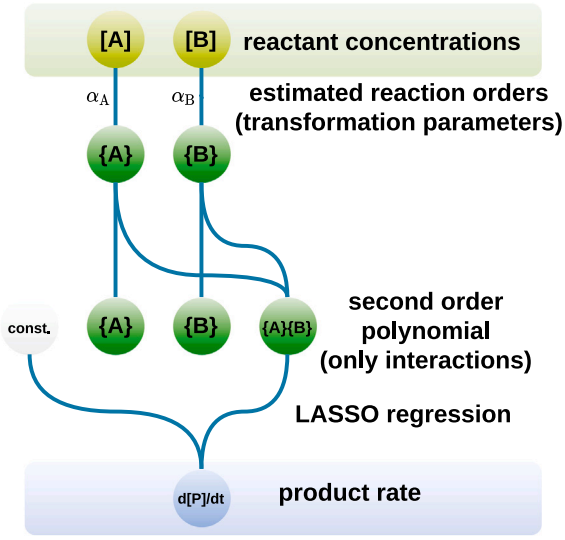

Fig. 2. Model structure for the relation between reaction factors and product rate. The factors (or input features), in this case, are the concentrations of the reactants. This is an example for a reaction with two reactants A and B and a single product P. The reactant concentrations are transformed by the current guess of their respective reaction orders (transformation parameters α). Based on these transformed variables, a polynomial model is constructed. In the regression, LASSO selects only those terms, which significantly contribute to the target variable, here the product rate.

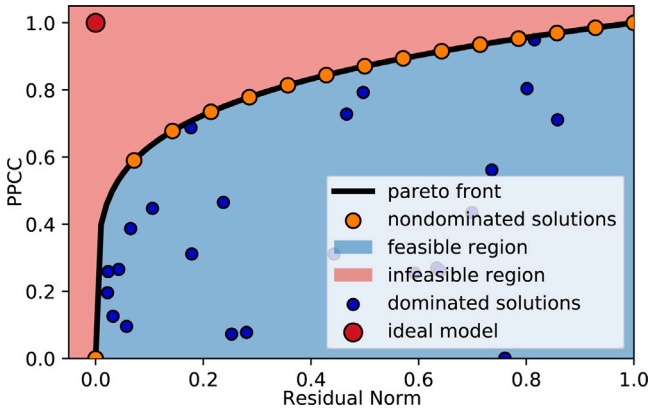

Fig. 3. Schematic of a pareto front. The PF separates the objective space into the so called feasible and infeasible regions. In our example, solutions located to the lower right of the PF are feasible, meaning there are parameter combinations, which realize such a objective value combination. Hypothetical solutions, located beyond the PF, are infeasible. Hence, points in this region of the objective space cannot be realized by any parameter combination. This is a result of the counteracting nature of the two objectives. In practice, the feasible region is often reduced to only the PF, as those points are the best possible feasible solutions. Within the feasible region, we distinguish between the points located on the PF and those somewhere below the front. The points on the front are referred to as non-dominated. In general, a solution A dominates another solution B, if A is not worse in any objective compared to B and superior to B in at least one objective. Therefore, the solutions on the PF are non-dominated, as there are no points, which are strictly better. On the other hand, all points within the feasible region but not on the PF are dominated.

errors into the model by choosing unreasonably complex polynomial terms to enforce small residuals. To achieve this, we explicitly consider the distribution of the model residuals in our cost function. As we assume the statistical noise in the underlying data to be normally distributed, we also want the residuals of our model to show such a distribution. A normally distributed model error implies that everything the model cannot explain is to be assumed as statistical noise. To quantify this criterion, the probability plot correlation coefficient (PPCC) introduced by Filliben [21] is used. The PPCC quantifies the normality of a statistical distribution as a value between 0 and 1. The closer to 1, the closer the distribution is to being normal.

We then end up with two separate objectives for the optimization, the quality of the fit and the normality of the error distribution. To avoid choosing one of these criteria over the other, we utilize the multi-objective evolutionary algorithm NSGA-II [17]. A multi-objective optimization usually does not result in one optimal set of parameters satisfying all objectives at the same time. Instead, we rather end up with a set of possible models forming the so called pareto front (PF) as illustrated in Fig. 3. These pareto-optimal models display different trade offs between the multiple objectives, where an increase in one objective would lead to a lower score in another. Hence, it is not straightforward to consider one pareto-optimal model to be better than any other one, as it is not possible to find a model with a higher score in all objectives.

In this context, the perfect model would achieve a PPCC value of 1 and a residual norm of 0. Looking at the schematic pareto front in Fig. 3, we see that these two objectives interfere with each other. The ideal model would be located in the upper left corner of the plot, within the infeasible region. Points on the pareto front, close to this ideal model, would provide a reasonable choice. Solutions with even lower residuals can only achieve this at the cost of neglecting the error distribution. Going to the other end of the pareto front, there are models which achieve an almost perfect normal distribution of errors but cannot provide a reasonable fit anymore. Ultimately, the solution offering the best trade off between all objectives needs to be chosen based on additional criteria and information. Below, we will discuss characteristic features of the pareto front, that can be taken as strong indicators for valid choices. A schematic, summarizing all relevant steps of our approach, is shown in Fig. 4.

3. Chemical processes

Three chemical processes are exemplarily studied in this work to illustrate the suggested algorithm. Synthetic concentration versus time data is obtained by numerically integrating the partial differential equations given by the reaction rate laws for these processes. A simplified one-dimensional transient plug flow reactor (TPFR) model is simulated, applying the method of lines [22] for the discretization along the reactor axis. Numeric approximations to the time derivatives of the concentrations have been obtained from central differences. For this purpose, the reaction has been simulated at different flow rates. In our idealized reactor model, a change in flow rate is equivalent to a change in residence time. At constant volume (reactor length), a change in residence time in a TPFR model is in turn equivalent to a change in reaction time (time before quenching) in a batch reactor. Therefore, in the limit of short residence times, we can estimate the reaction rate as the change of the product concentration at the reactor exit

$$\frac{d[\mathrm{P}]}{dt} \propto \frac{d[\mathrm{P}]_{exit}}{d\tau}$$
$$\frac{d[\mathrm{P}]_{exit}}{d\tau} \approx \frac{[\mathrm{P}]_{exit}(\tau+\delta\tau) - [\mathrm{P}]_{exit}(\tau-\delta\tau)}{2\delta\tau} \tag{6}$$

where t denotes the simulation time and τ refers to the residence time in the reactor. $\delta\tau$ is the finite change in residence time used to approximate the derivative. As we are not interested in the absolute values of the rates, in this work these derivatives with respect to the residence time τ will be used as approximations to the actual rate. The examples discussed in the following sections are all evaluated for short residence times and thus describe the low conversion limit of the respective chemical processes. In this limit, the overall concentrations are still very sensitive to the chemical reaction kinetics rather than to the transport through the reactor tube. After all, the effective models we construct are designed to only describe the reaction kinetic part of the reactor dynamics. Consequently, we generate training data for our algorithm in a way that is most sensitive to kinetic quantities.

The proposed algorithm is designed to operate on experimentally measured data. According to the central limit theorem [23], the accumulated effect of various noise sources in an experimental setup tends

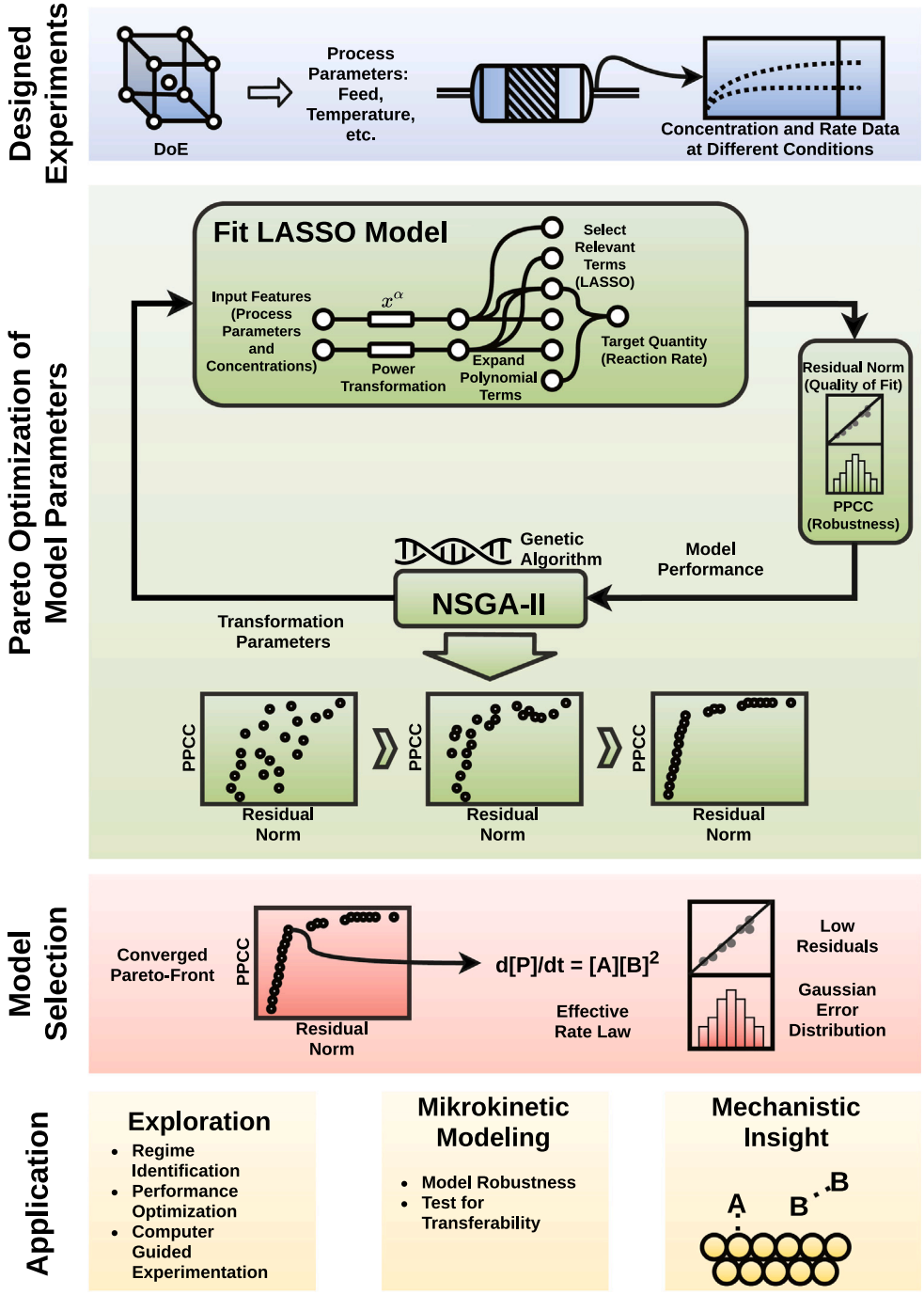

Fig. 4. Schematic of the sparse approximation approach. The proposed applications of the resulting kinetic models will be discussed in the following.

to result in a normally distributed measurement error. This gets more and more pronounced, the larger the data set is. As our algorithm relies on such normally distributed noise, we need to ensure that this distributional assumption is also valid for our artificial data sets. To simulate statistical noise of the experimental setup, a noise term based on normally distributed random numbers is therefore added to the computed rate data. The magnitude of this noise term is 1% of the mean value of the respective rate data set. Details on the influence of the magnitude of this noise term are shown in the SI. Measurement errors for reaction rates around 1% to 2% are encountered in literature [24,25] and a value of 1% lies well within the region, where the fitted rate models are not significantly affected by the noise.

The resulting data sets consist of the reactant concentrations at the reactor exit as well as the change in concentration or rate of the product species at different reaction conditions. In the examples shown in this work, these reaction conditions are characterized by differences in the feed composition (reactant concentrations) and temperature. It is important to properly sample these conditions, in order to be able to capture especially the interaction effects at play. For this reason we apply factorial experimental designs [14]. In these factorial designs, instead of keeping all but one factor fixed and vary each individually, all factors are varied at the same time. The simplest example of such a design is called a full factorial design (FFD). In a FFD, all possible combinations of factor settings are sampled. Again, we look at the example of the simple reaction A + B → P. If we want to vary the feed concentration of A and B, we first have to decide, how many settings we want to allow. Suppose we want to look at three different concentrations of both A and B, the corresponding FFD would dictate

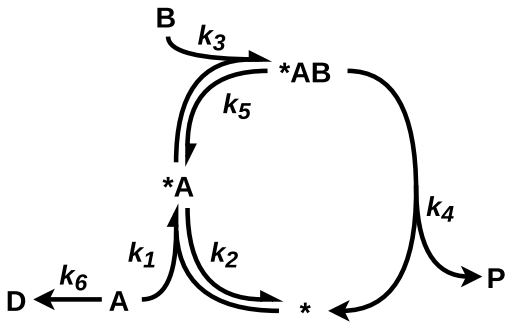

Fig. 5. Reaction network of the catalytic cycle model. Reactants A and B form an activated complex with a catalyst (*), which reacts to form the reaction product P in a subsequent step. The k_i are rate constants for the respective elementary steps.

$3 \cdot 3 = 9$ necessary experiments to sample all possible combinations. However, the amount of experiments for a FFD quickly becomes infeasible, if we increase the number of levels or factors. Therefore, such FFDs are only applicable for simple experimental setups or, in our case, simple reaction examples. The advantage of the FFD is that it contains the maximum amount of information possible about higher order and interaction effects. For more complex problems, we can often reduce the number of experiments by sacrificing information on effects beyond a certain order. Such designs are then generally referred to as fractional factorial designs. Most of the test cases in this study deal with simple reactions with only two to three factors. For these cases, we will use FFDs to sample the respective reaction conditions. In a later example, we will make use of a very efficient screening design, to reduce the number of required experiment for a slightly more complex chemical system.

3.1. Hydrogen bromine reaction

Hydrogen (H_2) and bromine (Br_2), in gas phase, can undergo a reaction to form hydrogenbromide (HBr),

$$H_2 + Br_2 \rightarrow 2\ HBr \quad . \tag{7}$$

This reaction is a classic example, where the apparent initial reaction orders are not obvious from the stoichiometric coefficients in the formal reaction equation. This indicates that a more complex microkinetic mechanism takes place. Bodenstein and Lind [26] described an effective rate law for the formation of HBr as

$$\frac{d[HBr]}{dt} \propto [H_2] \cdot [Br_2]^{1/2} \quad . \tag{8}$$

This law was found to be valid in the early stages of the reaction in a regime, where $[HBr] \ll [Br_2]$. It corresponds to an effective reaction order of 1.5 which could later be explained by the, now well established, radical reaction mechanism [27].

3.2. Catalytic cycle

The second reaction process consists of a simple model catalytic cycle. Two reactants A and B form a complex with the catalyst in a pre-equilibration step. A subsequent reaction forms the reaction product P and restores the catalyst's initial state. Reaction mechanisms of this type are frequently encountered in homogeneous catalysis and enzymatic reactions. The reaction equations are given in Fig. 5. This specific reaction model is based on work by Pollice [13]. We differentiate between two cases: the undisturbed catalytic reaction and the cycle with a side reaction ($k_6 \neq 0$) leading to a decomposition of reactant A.

Table 1
Model parameters and objective values for five selected pareto-optimal HBr models from the pareto front in Fig. 6.

	$\alpha_{[Br_2]}$	$\alpha_{[H_2]}$	Residual	PPCC	Dominating LASSO term
Solution 1	0.658	1.273	0.259	0.961	$\{Br_2\} \cdot \{H_2\}$
Solution 2	0.518	0.985	0.321	0.989	$\{Br_2\} \cdot \{H_2\}$
Solution 3	0.486	0.916	0.352	0.995	$\{Br_2\} \cdot \{H_2\}$
Solution 4	−4.066	2.908	1.258	0.995	$\{H_2\}$
Solution 5	−7.121	4.511	1.867	0.996	$\{H_2\}$

3.3. CO oxidation on Pt

The catalytic oxidation of carbon monoxide to carbon dioxide is a classical example for a reaction following a Langmuir–Hinshelwood mechanism [28]:

$$\begin{aligned} CO\ + *\ &\rightleftharpoons\ *CO \\ O_2 + 2\,* &\rightarrow 2\,*O \\ *O + *CO &\rightarrow CO_2 + 2\,* \end{aligned} \tag{9}$$

The oxidation of CO exhibits several interesting kinetic phenomena. For example, CO oxidation on a Pt catalyst is known for its oscillatory kinetics. The simple irreversible kinetic model used in this work is based on kinetic equations by Ertl and coworkers [29]. The detailed mechanism with all elementary steps and rate constants is included in the SI.

4. Results and discussion

4.1. Isothermal hydrogen bromine reaction

As a first test, we apply our algorithm to determine an effective rate law for the hydrogen bromine reaction at constant temperature. A set of concentration and rate data obtained at 25 different feed compositions is used to deduce the reaction orders in both reactants. This data set as well as the underlying data sets for the following sections are explicitly provided in the SI.

As we deal with two factors (the concentrations of H_2 and Br_2) in this first example, we choose a 2nd order ansatz for the isothermal HBr problem. Feeding reactant concentrations and product rates into our algorithm, we obtain the pareto front shown in Fig. 6(a).

Each point on the pareto front resembles a possible solution to the modeling problem. These solutions differ in the transformation parameters α on the reactant concentrations, i.e. the apparent partial reaction orders, and they differ in the set of dominant model terms selected by LASSO. Recall that these model terms are combinations of reactant concentrations which themselves are subject to a nonlinear transformation through the reaction order. For the isothermal HBr problem, the pareto front is composed of three distinct domains of solutions. The majority of solutions show comparably small residuals and are spread over a wide range of PPCC values. The other two domains show a slight increase in PPCC at the cost of a significant increase in model residuals. The transformation parameters, i.e. the reaction orders in the dominating regression term, for five representative solutions are given in Table 1. Solutions 1, 2 and 3 are located within the major domain of the front. All three models show reaction orders for Br_2 close to 0.5 and for H_2 roughly equal to 1. They thus reproduce the effective rate law from Bodenstein and Lind [26] with the underlying radical reaction mechanism. Solutions 4 and 5 are located in other domains of the front. They show vastly different parameter sets which cannot be reasonably interpreted as reaction orders.

Fig. 6(b) shows that this structure of the pareto front is also found in parameter space, where the distribution of possible reaction orders is arranged in three domains. This means, that all solutions within one domain of the pareto front show very similar transformation parameters. Switching from one domain to another does not only go with a

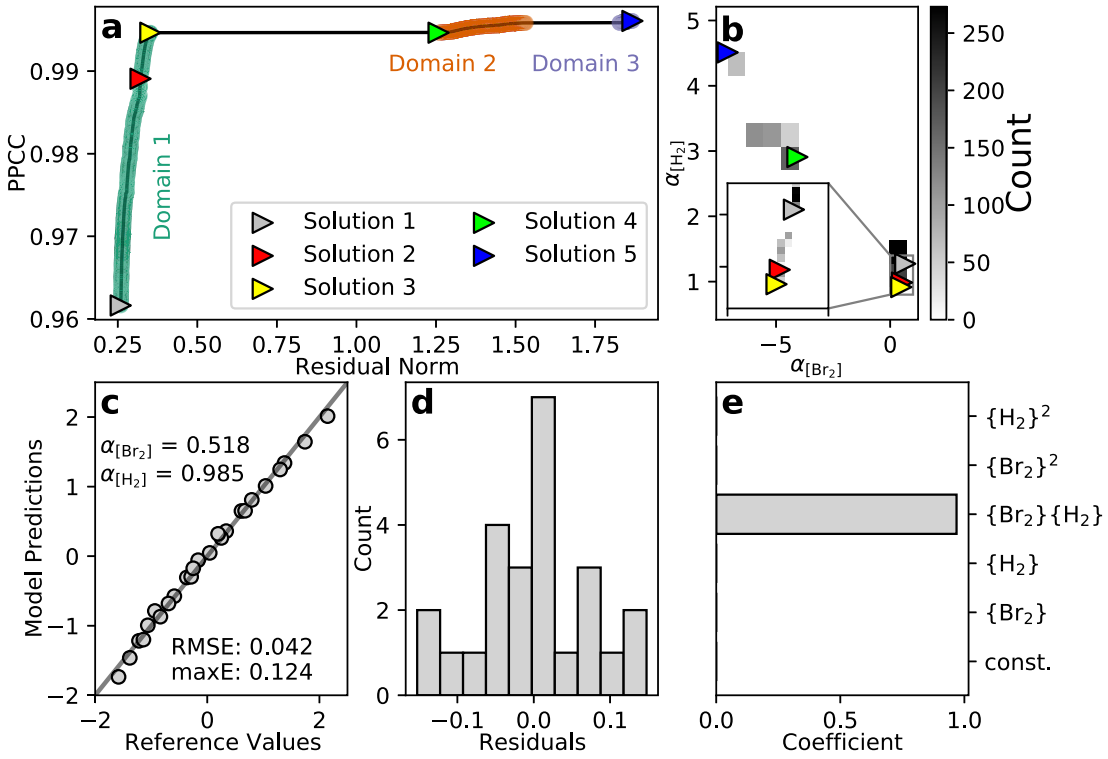

Fig. 6. (a) Pareto front for the isothermal HBr problem. The NSGA-II optimization is performed on a population of 1000 individuals. Five representative pareto-optimal solutions are highlighted with their respective effective reaction orders and model parameters given in Table 1. (b) Distribution of reaction orders α for both H_2 and Br_2 over the entire population. (c) Parity plot (predicted reaction rates vs. reference rates from microkinetic model) for solution 2. (d) Residual distribution for solution 2. (e) LASSO regression coefficients for solution 2.

change in parameters. Also the dominating LASSO coefficients change, meaning the model has an overall different functional form. For all solutions located in domain 1, only the interaction term $\{Br_2\} \cdot \{H_2\}$ contributes predominantly to the rate law, while the largest coefficient in solutions 4 and 5 corresponds only to the concentration of hydrogen $\{H_2\}$. This behavior exemplifies the flexibility of our approach. As soon as all possible trade-offs for a fixed functional form are exhausted, LASSO allows for the switching to an alternative model which can provide additional pareto-optimal solutions. This discontinuous structure of the pareto front also allows for a selection of interesting solutions without prior knowledge on the desired result. Pareto solutions located close to a discontinuity of the front resemble an extremum. These points describe either the maximum PPCC value, or the minimal residual norm for a given set of LASSO terms. Changing the transformation parameters (α) any further, would result in the LASSO optimization converging towards another set of terms and hence a different functional form of the resulting rate law. Further, the ideal solution should show a high PPCC value while maintaining a reasonably small residual norm. For most pareto fronts generated by our algorithm, these solutions are located in the upper left corner (e.g. in Fig. 6). For the HBr test case, the pareto front shows a clear kink in this region, followed by a large discontinuity. Following this reasoning for a chemically reasonable region of the pareto front, we should find the best possible solution at the upper edge of domain 1. And indeed solution 3 does show the expected reaction order of almost exactly 3/2. In the following sections we will show, that following such a heuristic does also result in chemically reasonable solutions for other systems, where we do not know the expected outcome in advance.

A closer look at the prediction quality of solution 2 is shown in Fig. 6(c). The partially negative values for the reaction rate in this figure result from the rescaling and shifting of the data during the modeling procedure. This is a technical requirement and details are given in the SI. Within the limits of 25 available data points, the error distribution for this model shown in Fig. 6(d) roughly resembles a Gaussian shape, resulting in its high PPCC value. While the LASSO regression is relatively robust against small sample sizes, evaluating this PPCC value will become the bottleneck of the proposed algorithm with respect to the required amount of data points. As the PPCC value ultimately compares two distributions, its significance depends on how well the shape of these distributions is captured by the available data. While the quality of the PPCC increases with larger data sets, it is generally capable of distinguishing between different distributions with remarkably small data, way below 100 points [21]. The 25 data points used in this example, however, are probably close to a minimum size in order to obtain reasonable results. A closer look at the convergence of the PPCC value with respect to the number of data points is provided in the SI. Further, the model error is small with a relative root mean squared error (RMSE) of 4.2%. Another interesting result is, that the LASSO algorithm chooses exactly only one polynomial term for this model of the HBr rate, perfectly reproducing the literature known effective kinetic law.

4.2. Temperature-dependent hydrogen bromine reaction

In a second data set, additionally also the temperature of the Hydrogen Bromine reaction is varied to illustrate the effect of temperature as another factor. A total of 45 concentration profiles were full factorially simulated within a temperature range between 840 K and 860 K. This temperature window is chosen to be that small, as the HBr reaction shows a very high sensitivity with respect to changes in temperature. Keeping the temperature variations small results in an effect on the reaction rate of similar magnitude compared to variations in feed concentrations. This prevents one factor from dominating the LASSO regression. Nevertheless, larger temperature ranges can also be realized within certain limits. A detailed discussion of this point, as well as the complete data set are provided in the SI. We can, however, not directly use the temperature data in order to determine its effect in our approach. If we assume an Arrhenius-like behavior for the

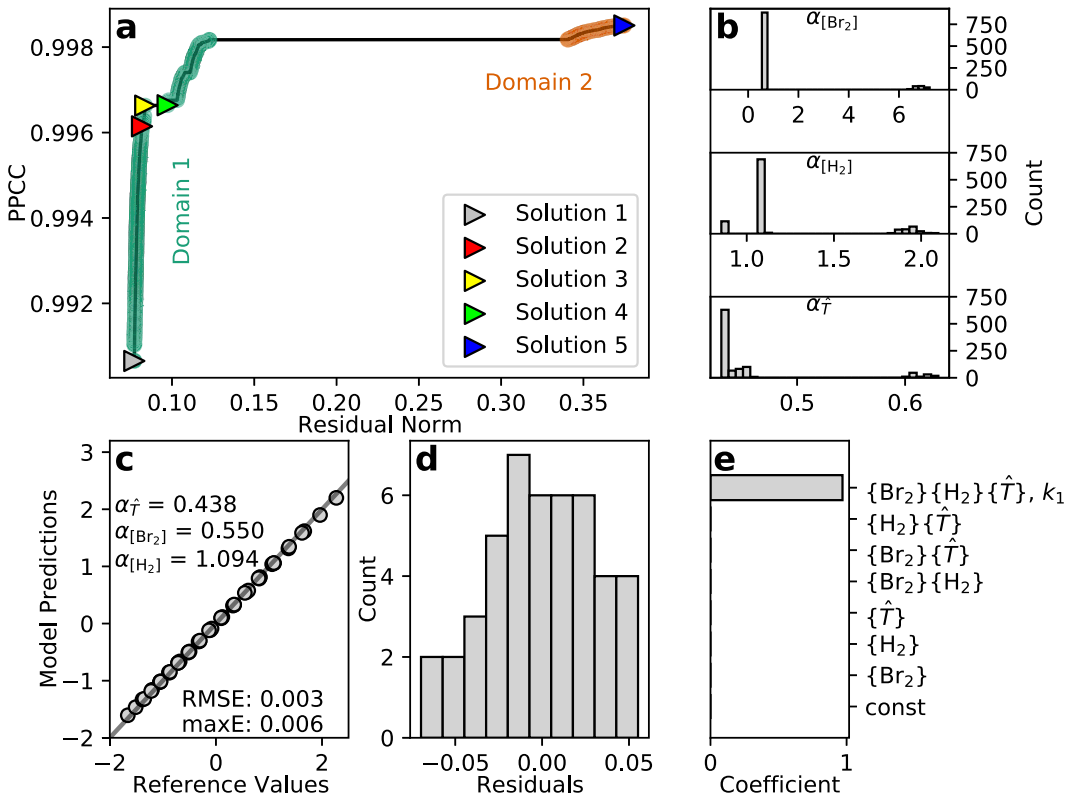

Fig. 7. (a) Pareto front for the temperature-dependent HBr problem. The NSGA-II optimization is performed on a population of 1000 individuals. Five representative pareto-optimal solutions are highlighted with their respective sensitivity parameters given in Table 2. (b) Distribution of reaction orders α for H_2, Br_2 and the scaled temperature values $\hat{T}$ over the entire population. (c) Parity plot for solution 2. (d) Residual distribution for solution 2. (e) LASSO regression coefficients for solution 2. k_1 indicates the regression coefficient, which significantly contributes to the solution. The resulting approximate rate law is given in Eq. (11).

temperature dependency of a general rate constant k, we can rearrange the corresponding Arrhenius equation as follows

$$\begin{aligned} k &= k_0 \cdot \exp\left(\frac{-E_A}{R \cdot T}\right) = k_0 \cdot \left(\exp\left(\frac{-1}{R \cdot T}\right)\right)^{E_A} \\ &= k_0 \cdot \left(\exp(-\beta_M)\right)^{E_A} = k_0 \cdot \hat{T}^{E_A} \\ \beta_M &= \frac{1}{RT}, \quad \hat{T} = \exp(-\beta_M) \end{aligned} \tag{10}$$

with k_0 the pre-exponential factor and R the universal gas constant and T the temperature. We define $\hat{T}$ as the rescaled temperature values $\exp(-\beta_M)$ and will use this notation in the following for brevity. By this rearrangement we can identify the apparent barrier E_A as the sensitivity parameter of the reaction rate to a change in $\hat{T}$. Thus, by feeding the algorithm these scaled temperatures additionally to the reactant concentrations, an estimate for the reaction barrier can be obtained. As we now have an additional input feature in terms of the $\hat{T}$ values, we choose a 3rd order polynomial ansatz for this temperature-dependent problem. Fig. 7 shows the resulting pareto front for this data set. The general structure of this front is similar to that of the isothermal problem.

Model parameters for five representative solutions are again given in Table 2. By means of the heuristics discussed for the isothermal case, solution 2 is located in the chemically reasonable area of the pareto front. Its reaction orders for hydrogen and bromine again reproduce the expected literature rate law of 3/2. The apparent activation barrier results in a sensitivity parameter of 0.438. Taking into account the proper rescaling of the $\hat{T}$ values, this inherently dimensionless sensitivity parameter directly translates into an activation barrier of 43.8 kcal/mol. Details on this scaling procedure are provided in the SI.

Levy [30] performed an experimental study on the high temperature kinetics of the HBr formation in the range between 600 K and 1400 K, extending the original study by Bodenstein and Lind [26] which was

Table 2
Model parameters and objective values for five selected temperature-dependent pareto optimal HBr models corresponding to the pareto front in Fig. 7.

	$\alpha_{[Br_2]}$	$\alpha_{[H_2]}$	$\alpha_{\hat{T}}$	Residual	PPCC
Solution 1	0.546	1.085	0.430	0.077	0.991
Solution 2	0.550	1.094	0.438	0.082	0.996
Solution 3	0.551	1.103	0.442	0.084	0.997
Solution 4	0.546	1.115	0.449	0.097	0.997
Solution 5	7.221	0.860	0.631	0.375	0.999

performed at around 500 K. The artificial data set created in this work was sampled around 800 K well within this high temperature regime. In experiment an apparent activation barrier of around 40.6 kcal/mol [30] was found, which is in reasonable agreement with the 43.8 kcal/mol determined by our algorithm. Looking at the resulting LASSO coefficients for this solution (Fig. 7(e)), now the 3-body interaction of all factors is dominating. For this solution the effective rate law would correspondingly have the form

$$\frac{d[\mathrm{HBr}]}{dt} \propto k_1 \cdot \left([\mathrm{Br_2}]^{0.550} \cdot [\mathrm{H_2}]^{1.094} \cdot \left(\exp(-\beta_M)\right)^{0.438}\right) \tag{11}$$

with k_1 the LASSO coefficient for the dominant term as indicated in Fig. 7(e).

4.3. Model catalytic cycle without side reaction

As a second prototypical case for reaction kinetics we now turn to a simple model catalytic cycle. One or more reactant molecules form an active complex with the catalyst in a pre-equilibrium. This active complex then irreversibly reacts to form the product in a rate-determining step, regenerating the catalyst. To test, whether our algorithm is able to capture the kinetics of such a catalytic process, we create test

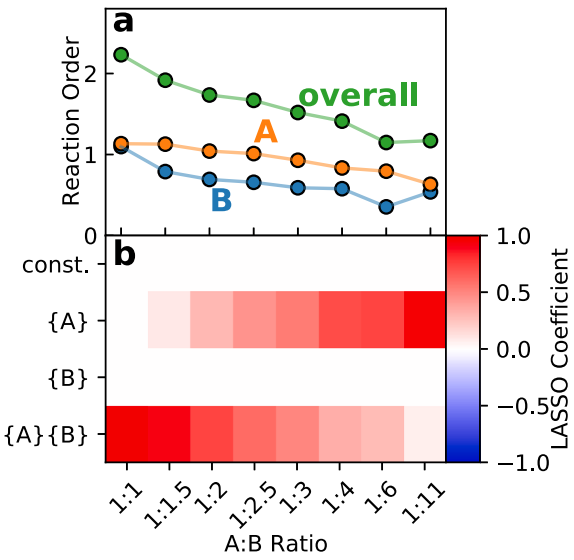

Fig. 8. (a) Effective reaction orders versus excess in reactant B for the model catalytic cycle without side reaction. (b) LASSO coefficients for the selected solutions at different A:B ratios. The coloring indicates the magnitude of the coefficients, which corresponds to the significance of this term to the overall solution. At a feed ratio of 1:1 only the interaction between A and B contributes to the corresponding rate law, while with increasing excess in B the linear term in A becomes more and more dominant.

data based on the model reaction network given in Fig. 5. The first test deals with the catalytic cycle without any side reaction ($k_6 = 0$). Concentration profiles are obtained for a range of different feed compositions starting at a 1:1 ratio of A to B up to a tenfold excess in B. For each feed ratio, 25 data points based on variations in the feed concentrations of A and B are simulated and an analysis using the modeling algorithm with a second-order ansatz for the considered LASSO terms is performed. For each analysis, one model is selected from the pareto front with the same heuristics regarding discontinuities in mind as discussed before for the HBr problem. The resulting reaction orders in A and B for these selected solutions are shown in Fig. 8.

While initially the orders in both reactants are close to 1 at a feed ratio of 1:1, the sensitivity of the reaction rate to the concentration of B quickly decreases to zero. The overall reaction order thus decreases from an initial second-order process to a first order for large excess values. This is to be expected, as ultimately the reaction rate only depends on the concentration of the limiting reactant. This simple example illustrates that the algorithm is capable of identifying such changes in the reaction regime, as clearly visible by the change in reaction orders. Further, as can be seen from Fig. 8(a), the sensitivities with respect to the individual reaction orders are observed individually. While the apparent order in reactant B steadily decreases with an increasing excess in B, the order in A only shows a slight decrease. The overall reaction order concomitantly switches from initially a second-order process depending on A and B to an effective first-order reaction limited only by A. This way it may easily be deduced which reactant is limiting the overall rate without any initial assumptions. Further insight can be provided by the LASSO regression coefficients, as illustrated in Fig. 8(b). With increasing excess, the dominating LASSO term switches from the product of both reactant concentrations to purely the concentration of A. This also explains the increasing deviations of the reaction order in B from the expected behavior, as the overall kinetic model becomes less and less sensitive to the concentration of reactant B. In the extreme case of an A:B ratio of 1:11, the term {B} does almost not contribute to the overall model anymore and hence also the reaction order in B is not well defined.

4.4. Model catalytic cycle with substrate decomposition

In a next step, we introduce the decomposition of reactant A as a side reaction. This leaves us with two target variables, the reaction rate towards product P and side product D. First, we set up a data set of 25 points varying only the feed concentrations at a constant temperature of 700 K. For the model ansatz again 2nd order terms are used, as we have two input features in the concentrations of A and B. It is straightforward to generalize our algorithm to identify surrogate rate laws for multiple targets. The main difference is an increase in the dimensionality of the objective space for the optimization, as now PPCC values for both $\frac{d[P]}{dt}$ and $\frac{d[D]}{dt}$ need to be considered. Consequently, this also entails a higher-dimensional pareto front, complicating the selection of a final solution. The resulting pareto front is shown in the SI and exhibits a similar structure as the previously analyzed 2D examples. Hence, at least for this 3D example, a manual model selection by visual inspection was still possible as before.

The optimum solution according to the heuristics discussed before yields the following rate laws for equal feed of A and B

$$\begin{aligned} \frac{d[P]}{dt} &\propto [A]^{1.03} \cdot [B]^{1.07} \\ \frac{d[D]}{dt} &\propto [A]^{1.03} \quad . \end{aligned} \tag{12}$$

Again, the algorithm accurately captures the expected behavior. The rate towards the reaction product P is identified to be dependent on both reactants A and B, while only A contributes to the formation of side product D. Additionally introducing temperature effects as done for the HBr example leads instead to complications. In principle, it would also be possible to obtain the two activation barriers for the individual reactions within one model by using the scaled temperatures as an input variable twice. However, to make sure, that only one temperature term contributes to each rate expression, such that the two barriers are disentangled, would result in a constrained LASSO [31] problem. Unfortunately, algorithms for the constrained problem are not yet commonly implemented in customary statistics packages. In fact, the same issue arises for the partial reaction orders of the reactants if the reaction network of interest contains different reaction pathways with drastically different kinetics. If, for example, the decomposition of A would be a second order process, e.g. due a self activation of the reactant being necessary, we would require two different transformation parameters for the concentration of A. In that case, as discussed before, the LASSO would need to be constrained to ensure disentangled results for the two reaction products. A workaround is to fit two independent models for both product species. We perform this individual analysis of the two target rates based on one data set. For this set we vary not only the feed concentrations but also the temperature to enable an estimation the activation barrier. Sampling these three factors in a 3-level FFD, we end up with 27 data points, cf. SI. Having one additional input feature, the rescaled temperatures, we increase the order of ansatz terms in this analysis to three. For both target variables, the rate towards P and the rate towards D, the algorithm produces a (now 2D) pareto front and a final solution can be selected as before. The resulting model parameters for these individual solutions are given in Table 3. The activation barrier for the product formation is identified to be 6.60 kcal/mol, corresponding to the barrier of the rate limiting decomposition of the activated catalyst complex. This is in nice agreement with the barrier for this step in the microkinetic model of 6.00 kcal/mol. Also for the side reaction, an activation barrier of 5.10 kcal/mol is obtained, in accordance with the reference value of 5.00 kcal/mol.

4.5. CO oxidation on Pt

In industrial applications, the question frequently arises under which conditions the regime of the reaction system may change. This is important for the optimization of process and reactor parameters. To

Table 3
Model parameters and objective values for selected models for the production rate of products P and D in the catalytic cycle with side reaction. The corresponding pareto fronts are given in the SI.

	$\alpha_{[A]}$	$\alpha_{[B]}$	$\alpha_{\hat{T}}$	Residual	PPCC
$\frac{d[P]}{dt}$	1.143	1.159	0.066[a]	0.159	0.995
$\frac{d[D]}{dt}$	1.131	−0.827[b]	0.051[a]	0.148	0.981

[a]In 100 kcal/mol.

[b]Not selected in LASSO solution.

this end, our robust modeling approach can provide valuable insight. As previously shown for the HBr and model catalytic cycle examples, the sensitivity parameters (apparent reaction orders) change from one regime to another. Additionally, however, also the LASSO terms of the overall model can vary. This gives the algorithm enough flexibility to only select the limiting terms for the kinetic model of each respective regime. For CO oxidation on Pt, we know that there is a drastic change in surface coverage at around 500 K from a partially CO covered to a purely oxygen covered catalyst [32–34]. To identify potential temperature regimes we need to determine effective kinetic models over a wide range of temperatures and analyze how the resulting models change qualitatively in various regimes. This comes with the advantage that we do not need any prior knowledge about the microkinetic mechanism underlying the observed rates. All we need to specify is the variable input factors and the range of parameter values we are interested in.

It is especially important though to properly design the experiments for such a scan over a wider temperature range, as the goal is to identify relevant regimes with as few data points as possible. As an example, we generate kinetic data for the CO oxidation between 300 K and 700 K based on experimental designs [35] translated along the temperature axis. Details on the construction of these designs are given in the SI. Using this data set we can cut out smaller temperature ranges and fit local surrogate models. For reference, we create a large data set based on a FFD at every temperature with a total of 576 data points. Smaller partially overlapping subsets of 108 data points each, are cut out to determine local models for every temperature range. These local models are constructed using the proposed algorithm. An expansion of the polynomial terms up to 3rd order is used, as we need to take into account four input features, the concentrations of three species (CO, O_2 and CO_2), as well as temperature. Such a procedure results in a large number of pareto fronts, from which one solution needs to be selected respectively. This mandates an automated way of selecting the ideal solution to reduce on the one hand the required manual work, but on the other hand also to ensure more comparable and bias-free results. Here, we employ a rudimentary heuristic based on the same discontinuity arguments as before, which would guide a manual selection process. Details regarding this heuristic are explained in the SI. By observing the LASSO coefficients of these selected solutions we obtain a map, showing which model contributions are dominant at a given temperature. Fig. 9 shows temperature maps for the CO oxidation on Pt. The resulting map based on the full factorial reference is given in Fig. 9(c). It can be clearly seen, that the dominant LASSO terms undergo a drastic change over the range from 300 K to 700 K. Accordingly, we can roughly partition this map into four regimes and link these to characteristic changes in the underlying microkinetic model. To rationalize these findings, in Fig. 9(a) we show the temperature dependence of the dominating rate constants from the microkinetic model. These rate constants correspond to elementary processes of the catalytic process, namely the adsorption of O and CO (ads_O, ads_CO), the desorption of CO (des_CO) as well as the surface reaction between the adsorbed species (reac). As further reference Fig. 9(b) shows the concentration of adsorbed CO, resolved along both the temperature and the reactor axis.

Regime 1 (300 K to 420 K). At low temperatures up to about 420 K the interaction term $\{\hat{T}\}\{CO\}$ is dominating the models. This corresponds to the interaction of the temperature effect and the concentration of CO. Looking at the behavior of the rate constants of the elementary processes in Fig. 9(a), we see that in this same low temperature range, the rate constant of the surface reaction between the adsorbed species (reac) changes dramatically, explaining the significant sensitivity to temperature. Furthermore, over the entire temperature range, the adsorption of oxygen is at least three times faster compared to the adsorption of CO, making CO the limiting species for the formation of carbon dioxide.

Regime 2 (420 K to 540 K). In the subsequent transition region from 420 K to 540 K the temperature influence is less pronounced. The dominating terms now contain mostly the concentrations of oxygen ($\{O_2\}$) and CO ($\{CO\}$). Looking at the rate constants (Fig. 9(a)) there is no significant change to observe in this temperature range. The CO coverage (Fig. 9(b)), on the other hand, undergoes a drastic change in this regime, from a partially CO covered surface to an approximately CO free catalyst above 530 K. For this reason, the relative contributions of oxygen and CO to the overall effective kinetic behavior do heavily depend on the temperature. This may explain, why the algorithm is not able to reduce the effective kinetics to a single dominating term for this range of temperatures.

Regime 3: (540 K to 600 K). The third regime again shows a stronger temperature dependence. The three body interaction ($\{\hat{T}\}\{O_2\}\{CO\}$) is the dominant term in this region. A similar reasoning as in regime 1 also applies here. The drastic change in the relative rate of CO adsorption and desorption leads to the observed temperature dependence.

Regime 4: (600 K to 700 K). The high temperature regime above 600 K is again dominated by effects of reactant concentrations ($\{O_2\}\{CO\}$). Temperature contributions are not significant and the rate is basically only limited by the small fraction of adsorbed CO on the mostly oxygen covered surface (c.f. Fig. 9(b)). This coverage of CO highly depends on both the gas phase concentrations of CO and oxygen, as in the underlying microkinetic model both species compete for the same adsorption sites.

In reality it is, of course, not possible to experimentally measure almost 600 concentration profiles for such an analysis. Therefore, we require a means of more efficiently sampling the relevant data. In Fig. 9(d) we show the temperature map obtained based on a combination of low-resolution screening designs with a total of only 90 profiles. The corresponding subsets consist of only 15 data points at three temperatures each. This extremely small amount of data challenges the estimation of the PPCC values. Still, as we are mostly interested in a rough estimate of the change in LASSO coefficients, such screening designs can be applied. In order to fit an accurate surrogate model, applicable for the prediction of reaction rates, a larger data set will be required. Looking at Fig. 9(d), we can see that the qualitative features of the map are very similar to the full factorial reference in Fig. 9(c). Regimes 1 (300 K–420 K) and 2 (420 K–540 K) are well captured by the screening designs, and also the high-temperature behavior is reproduced. The main difference between the two designs occurs in regime 3 (540 K–600 K). Here, the temperature contribution is missing and the dominant term is $\{O_2\}\{CO\}$ instead of $\{\hat{T}\}\{O_2\}\{CO\}$. Still, there is a small contribution by $\{\hat{T}\}\{O_2\}\{CO\}$ also present for the screening design. This shows, that an initial assessment of relevant temperature regimes can already be obtained with a much smaller amount of data, while a detailed modeling of the most interesting regimes could require additional sampling in a second step.

Assuming that we do not have access to the extensive full factorial data, we would identify three temperature regimes based on the temperature screening map. In addition to that, we obtain a rough estimate on how the reaction rate behaves with changing temperature (blue crosses in Fig. 9(e)). It turns out, that the reaction rate reaches

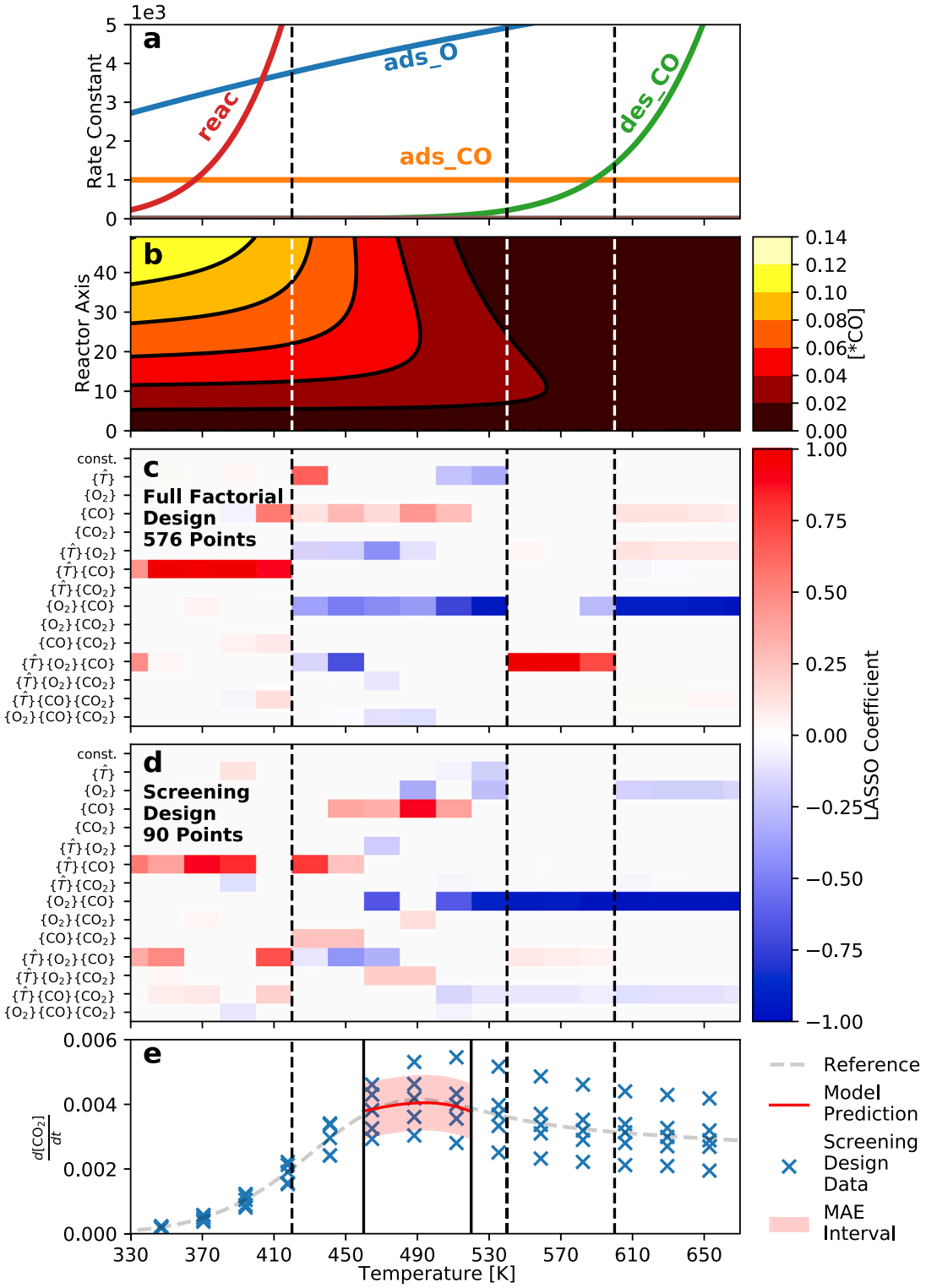

Fig. 9. Temperature scans for the CO oxidation model. (a) Temperature dependence of the dominant rate constants from the underlying kinetic model (Details provided in the SI). (b) Concentration of adsorbed CO species on the catalyst sites versus temperature from the microkinetic model. While there is some CO coverage present at lower temperatures, above 570 K there is basically no adsorbed CO left. (c) Reference Temperature map based on a full factorial experimental design with a total of 576 points. (d) Temperature map based on a set of screening designs with a total of 90 data points. (e) Rate data versus temperature for the entire range. The blue markers refer to all simulated rates at different conditions within the 90 point screening design. The gray dashed line gives the ideal rate at constant feed composition ([CO] = [O_2] = [CO_2] = 1 mol/l) and acts a guide to the eye. The red line and area shows the fitted model prediction in the regime between 460 K and 520 K with the mean absolute error (MAE) interval.

its maximum in the range around 500 K. Looking at the temperature map, this maximum region is located within regime 2. As mentioned before, regime 2 does not show a clear dominant model term, which can be attributed to the constant change in coverage in this temperature range. The maximum in reaction rate, however, can further explain the appearance of multiple model terms in the temperature map. To form such an extremal point in terms of the rate, at least two counteracting processes need to take place in parallel. These processes are represented

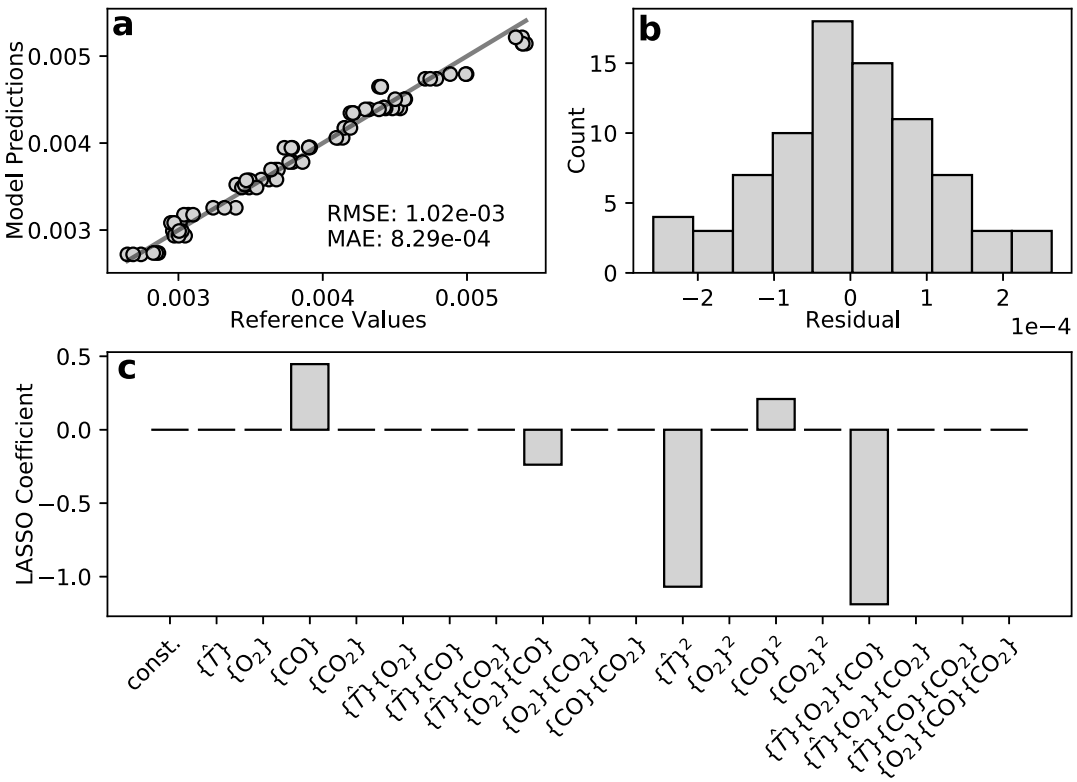

Fig. 10. Details for the maximum region (460 K to 520 K) effective kinetic model identified by the screening approach of Fig. 9(b): (a) Parity plot of the model prediction versus the reference from the microkinetic model. (b) Residual distribution. (c) LASSO coefficients for the maximum region model.

by different LASSO terms with their relative importance determined by the coverages of the reactants. In order to more accurately model this region of interest, additional data points are acquired in the range between 460 K and 520 K based on a 3-level full factorial design yielding 81 additional points. From the screening design data we know that the maximum rate should be located in this region. Using this more extensive data set we can fit a surrogate model for this entire temperature range. In this model, also quadratic terms are included to better describe the extremum. We have discussed before, that purely quadratic and higher power terms are usually neglected in our model ansatz. However, if we know that there is an extremum located within our region of interest, including quadratic terms can be beneficial. The red line in Fig. 9(e) indicates the model prediction for the reaction rate at constant feed composition ([CO] = [O_2] = [CO_2] = 1 mol/l) with an area showing the mean absolute error (MAE) of the model based on the training data. Using the additional data points, our algorithm is able to accurately reproduce the curvature of the reaction rate as a function of temperature in this region. The reference data (gray dashed line) for the rate at the same constant feed composition falls well within the error interval. The model predictions here are only given for the region of interest between 460 K and 520 K. An extrapolation to both higher and lower temperatures is possible up to a certain degree. However, while polynomial models excel at locally approximating complex functions, one has to be careful with extrapolations beyond the training region. Extrapolation based on the presented model is discussed further in the SI. In Fig. 10 additional details on the fitted kinetic model for the maximum rate region are given.

Inspecting the LASSO coefficients (e.g. the selected model terms) in Fig. 10(c), we see that in nice agreement with the temperature map, multiple terms contribute to the final model. The two largest contributions come from the 3-body interaction ({$\hat{T}$}{O_2}{CO}) and the quadratic term ({$\hat{T}$}^{2}). The former indicates a strong interplay between CO and O_2 with a significant temperature dependence. Coming back to the strong changes in surface coverages in this temperature region, this is to be expected. The large negative coefficient of the quadratic temperature term on the other hand is in nice accordance with the maximum in rate observed in Fig. 9(e). The remaining model terms all depend on the concentration of CO ({CO}), with two of them ({CO} and {CO}^{2}) showing a positive coefficient. This points at the growing dependence of the overall rate on the concentration of CO at higher temperatures as also observed for the high temperature regimes in the temperature map. Finally, we need to analyze the sensitivity parameters (apparent reaction orders and barrier) determined by the algorithm (Table 4). The apparent activation barrier of around 7.600 kcal/mol is in reasonable agreement with the barrier for the surface reaction in the microkinetic model of 10.000 kcal/mol. A negative reaction order for O_2 indicates a detrimental effect of an increase in oxygen concentration on the reaction rate. We can understand this, as the underlying microkinetic model does only consider one type of adsorption site. Therefore, an increase of oxygen concentration may lead to a reduced adsorption rate of CO. At the same time, an increase in CO concentration will lead to an increase in reaction rate causing the positive reaction order in CO. The concentration of CO_2 does not influence the reaction rate, as no readsorption of the product species is considered. This leads to a reaction order of effectively zero in the concentration of CO_2. This detailed analysis of the maximum rate region (460 K to 520 K) shows, that our algorithm can not only be used to identify interesting kinetic regimes of an unknown system with very limited data. It can also reproduce the effective kinetics of a local regime in a subsequent step. Further, we can interpret the resulting

Table 4
Model parameters and objective values for the selected solution for the maximum region of the CO oxidation example. The corresponding pareto front is given in the SI.

$\alpha_{\hat{T}}$	$\alpha_{[O_2]}$	$\alpha_{[CO]}$	$\alpha_{[CO_2]}$	Residual	PPCC
0.760[a]	−0.610	0.306	0.004	1.346	0.997

[a]In 10 kcal/mol.

model function in a physically meaningful way. Of course, within this study, the physical meaning of the obtained models is limited by the assumptions of the underlying simulated reaction models.

Going beyond the pure interpretation of the model parameters and terms, we can also use the obtained models in an optimization process to identify reaction conditions which maximize the rate. For more complex reaction networks such an optimization could even target other properties like selectivities towards certain products. A closer look at the predictions along cuts through all dimensions of the rate model (the three feed concentrations and temperature) shows that our model does indeed not only describe the temperature trend but also the effect of the feed composition and could therefore be used in such an optimization. The corresponding contour plots are provided in the SI.

5. Conclusions

We presented a novel approach to the analysis of reaction kinetic data with the goal of not finding the most accurate model, but the most robust one. By analyzing both the residual norm and the distribution of errors for potential candidate solutions, we propose an algorithm to identify the optimal trade-off. With the main goal of reducing the amount of systematic error through model assumptions, we let the algorithm choose polynomial terms making use of the sparsifying LASSO regression method. Additionally, the algorithm determines sensitivity parameters which can be interpreted as physical quantities like activation barriers or apparent reaction orders. Besides yielding rate law like models, the algorithm can further provide valuable insight into the underlying processes. This way, a meaningful analysis of reaction kinetics can be performed without knowledge of the exact reaction mechanism. Emphasizing the aspect of robustness in such an analysis may become especially relevant with respect to transferability of the obtained models.

We showed that such a procedure reproduces well understood reaction systems like the hydrogen bromine reaction or the catalytic model cycle without prior knowledge. In the HBr case, the empirically observed fractional reaction order of 3/2 was automatically detected by our method. At the same time we obtained an interpretable rate law capturing the effect of not only the reactant concentration but also the influence of temperature. The catalytic cycle presented in Section 4.4 was equally well described and we were able to capture the change in effective kinetics with an increasing excess of one reactant by monitoring the sensitivity parameters resulting from the automatic analysis. Going one step further towards industrially relevant processes, an analysis of the CO oxidation kinetics on Pt over a wide temperature range provided both, information about process regimes, as well as an accurate surrogate model for the region of interest. While these surrogate models cannot provide detailed knowledge about the reaction mechanism, valuable indications can be gathered by a thorough analysis of both the transformation parameters (α) and the polynomial terms selected by LASSO. This way we can learn how reaction orders and limiting species may change with reaction conditions like the feed composition.

It is important to mention, that this entire procedure relies on rationally designed experimental data sets, something not frequently encountered even in recent studies. A factorial sampling of experimental data, varying multiple factors at the same time is vital for a meaningful inference on interaction terms. As known from chemical reaction kinetics and also shown by our results, these interactions terms are most often the key contributions to the rate laws of various reactions. The artificial data sets used in this work rely on classical experimental designs. For the CO oxidation example, making use of a simple construction based on established screening designs we were able to achieve accurate results in a data efficient way. In the future, the data requirements for such an analysis could even further decrease by employing especially designed optimal designs [36,37] for such specific use cases. Applying this algorithm to more complex reaction systems and especially working on real experimental data will be necessary to benchmark against established methodologies. Experimentally obtaining suitable training data for such an analysis will be a challenge in its own right. The algorithm requires kinetic data reflecting the chemical reaction kinetics. Therefore, such experiments need to be performed in specialized reaction sensitive reactor setups [38–40] which are not dominated by transport phenomena. While a polynomial approximation of the data in a transport limited reactor could in principle yield accurate results, any kind of interpretability in terms of kinetic parameters is lost in such cases. Provided suitable training data, however, our approach promises to provide efficient kinetic models which, coupled to varying transport models, should even be transferable between different reactor geometries. This would allow for a straightforward upscaling going from idealized lab scale experiments to real industrial reactors. Obviously, our method is not designed to describe the global kinetics of a system over a wide range of conditions. The resulting models are also not intended to describe transport phenomena in chemical reactors. This approach rather resembles a semi local sensitivity analysis, which provides simple and well behaved kinetic models.

The formulation of this method allows for a straightforward extension in terms of both input and response variables to deal with more extensive reaction networks. In addition to that, also the optimization cost function may easily be modified to employ other error models for specific use cases. In this context, it may also be useful to apply our algorithm to data from complex microkinetic models. Comparing the result of such an analysis on experimental data versus model predictions, both in terms of sensitivity parameters and dominant LASSO terms, could be an extremely sensitive measure of the quality of a microkinetic model. Further, the capability of our approach to obtain robust effective kinetic models without any prior knowledge of the system opens up a wide field of possibilities towards computer guided experimentation. By optimizing desired properties like turn over frequencies or selectivities on the level of the cheap effective kinetic models, an algorithm can propose potentially interesting parameter ranges to sample next in an iterative procedure. This way, a fully automated exploration of relevant regimes as well as optimization of the process conditions could be realized.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Funding: This work was supported by the Munich Catalysis-Alliance (MuniCat) of Clariant and TUM, Germany.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cej.2021.134121.

References

[1] M. Feinberg, Foundations of chemical reaction network theory, Springer International Publishing, 2019.
[2] N.G.V. Kampen, Stochastic processes in physics and chemistry, Elsevier Science, 1992.
[3] F.A. Lindemann, A. Svante, L. Irving, N.R. Dhar, J. Perrin, W.C.M. Lewis, Discussion on 'the radiation theory of chemical action', Trans. Faraday Soc. 17 (1922) 598, http://dx.doi.org/10.1039/TF9221700598.
[4] S. Per, Microkinetic simulation of catalytic reactions, Prog. Surf. Sci. 65 (2000) 65, http://dx.doi.org/10.1016/S0079-6816(00)00019-8.
[5] B.W.J. Chen, X. Lang, M. Manos, Computational methods in heterogeneous catalysis, Chem. Rev. 121 (2021) 1007, http://dx.doi.org/10.1021/acs.chemrev.0c01060.

[6] A. Mie, P. Chiara, R. Karsten, A practical guide to surface kinetic Monte Carlo simulations, Front. Chem. 7 (2019) 202, http://dx.doi.org/10.3389/fchem.2019.00202.
[7] G. Hendrik, M. Lubow, A. Sofia, T. Steffen, D. Olaf, Carmen: An improved computer-aided method for developing catalytic reaction mechanisms, Catalysts 9 (2019) 227, http://dx.doi.org/10.3390/catal9030227.
[8] D.G. Blackmond, Reaction progress kinetic analysis: A powerful methodology for mechanistic studies of complex catalytic reactions, Angew. Chem. Int. Ed. 44 (2005) 4302, http://dx.doi.org/10.1002/anie.200462544.
[9] B. Jordi, What is the order of a reaction?, Top. Catal. 60 (2017) 631, http://dx.doi.org/10.1007/s11244-017-0735-y.
[10] B. Jordi, A simple graphical method to determine the order in catalyst, Angew. Chem. Int. Ed. 55 (2016) 2028, http://dx.doi.org/10.1002/anie.201508983.
[11] B. Jordi, Variable time normalization analysis: General graphical elucidation of reaction orders from concentration profiles, Angew. Chem. Int. Ed. 55 (2016) 16084, http://dx.doi.org/10.1002/anie.201609757.
[12] G.E. Garrett, M.S. Taylor, A nonlinear ordinary differential equation for generating graphical rate equations from concentration versus time data, Top. Catal. 60 (2017) 554, http://dx.doi.org/10.1007/s11244-017-0739-7.
[13] P. Robert, A general fitting function to estimate apparent reaction orders of kinetic profiles, ChemRxiv (2019) http://dx.doi.org/10.26434/chemrxiv.7885760.v1.
[14] G.E.P. Box, N.R. Draper, Response Surfaces, Mixtures, and Ridge Analyses, Wiley, 2007.
[15] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. Soc. Ser. B 26 (1964) 211, http://dx.doi.org/10.1111/j.2517-6161.1964.tb00553.x.
[16] J.W. Tukey, Exploratory Data Analysis, Bd. 2, Addison-Wesley Publishing Company, 1977.
[17] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182, http://dx.doi.org/10.1109/4235.996017.
[18] S. Fadil, W.W. Symes, Linear inversion of band-limited reflection seismograms, SIAM J. Sci. Comput. 7 (1986) 1307, http://dx.doi.org/10.1137/0907087.
[19] T. Robert, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B 58 (1996) 267, http://dx.doi.org/10.1111/j.2517-6161.1996.tb02080.x.
[20] K.J. Laidler, A glossary of terms used in chemical kinetics, including reaction dynamics (IUPAC recommendations 1996), Pure Appl. Chem. 68 (1996) 149, http://dx.doi.org/10.1351/pac199668010149.
[21] J.J. Filliben, The probability plot correlation coefficient test for normality, Technometrics 17 (1975) 111, http://dx.doi.org/10.1080/00401706.1975.10489279.
[22] W. Schiesser, The Numerical Method of Lines: Integration of Partial Differential Equations, Academic Press, 1991.
[23] G.E.P. Box, W.H. Hunter, S. Hunter, Statistics for Experimenters, John Wiley sons New York, 1978.
[24] E.M. Cordos, S.R. Crouch, H.V. Malmstadt, Automatic digital readout system for reaction-rate methods, Anal. Chem. 40 (1968) 1812, http://dx.doi.org/10.1021/ac60268a018.
[25] J. L. Lane, D.R. McGregor, J. Mark Gezer, A. Aydin, The effect of experimental measurement uncertainties on rate modeling and reactor simulation, Chem. Eng. Commun. 76 (1989) 41, http://dx.doi.org/10.1080/00986448908940317.
[26] B. Max, S.C. Lind, Geschwindigkeit der bildung des bromwasserstoffs aus seinen elementen, Z. Phys. Chem. 57U (1907) 168, http://dx.doi.org/10.1515/zpch-1907-5711.
[27] K.F. Herzfeld, Zur theorie der reaktionsgeschwindigkeiten in gasen, Ann. Physics 364 (1919) 635, http://dx.doi.org/10.1002/andp.19193641504.
[28] T. Engel, G. Ertl, Elementary steps in the catalytic oxidation of carbon monoxide on platinum metals, Adv. Catal. 28 (1979) 1, http://dx.doi.org/10.1016/S0360-0564(08)60133-9.
[29] I. Ronald, E. Gerhard, Oscillatory kinetics in heterogeneous catalysis, Chem. Rev. 95 (1995) 697, http://dx.doi.org/10.1021/cr00035a012.
[30] L. Arthur, The high-temperature kinetics of the hydrogen-bromine reaction, J. Phys. Chem. 62 (1958) 570, http://dx.doi.org/10.1021/j150563a013.
[31] B.R. Gaines, K. Juhyun, Z. Hua, Algorithms for fitting the constrained lasso, J. Comput. Graph. Stat. 27 (2018) 861, http://dx.doi.org/10.1080/10618600.2018.1473777.
[32] R. Kissel-Osterrieder, F. Behrendt, J. Warnatz, U. Metka, H.-R. Volpp, J. Wolfrum, Experimental and theoretical investigation of CO oxidation on platinum: Bridging the pressure and materials gap, Proc. Combust. Inst. 28 (2000) 1341, http://dx.doi.org/10.1016/S0082-0784(00)80348-3.
[33] Y.Y. Yeo, L. Vattuone, D.A. King, Calorimetric heats for CO and oxygen adsorption and for the catalytic CO oxidation reaction on pt111, J. Chem. Phys. 106 (1997) 392, http://dx.doi.org/10.1063/1.473203.
[34] G. G. T. Kasun Kalhara, S. Mark, Co adsorption on pt(111): From isolated molecules to ordered high-coverage structures, ACS Catal. 8 (2018) 10225, http://dx.doi.org/10.1021/acscatal.8b02371.
[35] R.L. Placket, J.P. Burman, The design of optimum multifactorial experiments, Biometrika 33 (1946) 305, http://dx.doi.org/10.1093/biomet/33.4.305.
[36] S. Izabella, V. Ludvig, H. Gustaf, J. Erik, V. Conny, T. Johan, Anal. Chem. 89 (2017) 6491, http://dx.doi.org/10.1021/acs.analchem.7b00506.
[37] H. Radoslav, F. Lenka, R. Peter, A randomized exchange algorithm for computing optimal approximate designs of experiments, J. Amer. Statist. Assoc. 115 (2020) 348, http://dx.doi.org/10.1080/01621459.2018.1546588.
[38] J.M. Berty, Testing commercial catalysts in recycle reactors, Catal. Rev. Sci. Eng. 20 (1979) 75, http://dx.doi.org/10.1080/03602457908065106.
[39] L. Caldwell, An improved internal gas recirculation reactor for catalytic studies, Appl. Catal. 8 (1983) 199, http://dx.doi.org/10.1016/0166-9834(83)80080-3.
[40] E.M. Calverley, E.L. Lee, Y. De-Wei, T.J. Parsons, A small, well-mixed reactor for high throughput study of commercial catalyst pills, Chem. Eng. Sci. 151 (2016) 130, http://dx.doi.org/10.1016/j.ces.2016.05.011.