Kinetic models parameterised by \textit{ab-initio} calculations have led to significant improvements in understanding chemical reactions in heterogeneous catalysis. These studies have been facilitated by implementations which determine steady-state coverages and rates of mean-field micro-kinetic models. As implemented in the open-source kinetic modelling program, CatMAP, the conventional solution strategy is to use a root-finding algorithm to determine the coverage of all intermediates through the steady-state expressions, constraining all coverages to be non-negative and to properly sum to unity. Though intuitive, this root-finding strategy causes issues with convergence to solution due to these imposed constraints. In this work, we avoid explicitly imposing these constraints, solving the mean-field steady-state micro-kinetic model in the space of \textit{number of sites} instead of solving it in the space of \textit{coverages}. We transform the constrained root-finding problem to an unconstrained least-squares minimisation problem, leading to significantly improved convergence in solving micro-kinetic models and thus enabling the efficient study of more complex catalytic reactions.

\textbf{Keywords:} Micro-kinetic modelling, heterogeneous catalysis, surface chemistry, reaction networks
We have developed a constraint-free solver for determining surface coverages from steady-state mean-field micro-kinetic models, which are a crucial component for developing a mechanistic understanding of heterogeneous catalytic processes. This implementation leads to a more stable convergence in the iterative solution of such models, as depicted by the green symbols, while an approach employing constraints on coverages, as indicated by the red symbols, can fail to find the solution.

1 INTRODUCTION

Kinetic modelling is a fundamental tool for understanding chemical reactions in heterogeneous catalysis.\textsuperscript{1–3} Mechanistic analysis through kinetic models provides knowledge of the rate-limiting step of a reaction, which is vital to its optimization towards products of interest. Often, kinetic models are based on first principle calculations of activation and adsorption (free) energies, usually through Density Functional Theory (DFT). The resulting micro-kinetic models enable an atomistic understanding of catalytic processes.\textsuperscript{4}

In the past few years, so-called descriptor-based analysis\textsuperscript{5} coupled with micro-kinetic models have led to a systematic understanding of activity and selectivity trends in heterogeneous catalysis.\textsuperscript{6} This analysis is performed by constructing linear-scaling relations,\textsuperscript{7} in which the adsorption and activation energies of and between all intermediates in a reaction network (typically of the order of about 20-100 species) are linearly mapped onto the adsorption energies of a handful (about 2-3) reaction intermediates. This reduction in the problem’s dimensionality allows for an effortless estimation of reaction rates based on a few computed adsorption energies. Subsequently, the highest reactivity and, therefore, the most suitable catalyst is readily determined based on the optimal values of the descriptors.
Several implementations have greatly facilitated the applications of kinetic modelling to reactions in heterogeneous catalysis.\textsuperscript{5,8–11} One such implementation specific to descriptor-based analysis is the Python-based Catalysis Microkinetic Analysis Package (CatMAP).\textsuperscript{5} CatMAP solves a mean-field micro-kinetic model to determine the coverages of all intermediates and hence the rates of the involved elementary reactions. The software determines the steady-state solution based on a set of coupled non-linear algebraic equations. One of the main benefits of using CatMAP is that it removes the need for individual implementations in kinetic modelling studies and provides a platform for automating complex reaction networks. It has enjoyed wide use in kinetic modelling and heterogeneous catalysis research.\textsuperscript{12–15}

For highly complex reaction networks, the root-finding strategy in CatMAP struggles to determine the steady-state coverage yielding incomplete solutions or long convergence times. This difficulty in convergence is partly due to the constrained nature of the root-finding algorithm used to determine the steady-state coverage. The solution must obey two constraints. First, all coverages must be non-negative. Second, the sum of coverages of all species must be precisely one, i.e. a species can, at most, occupy the total number of available surface sites. We eliminate both these constraints by mathematically transforming the coverage degrees of freedom. This transformation results in an unconstrained root-finding approach, leading to smoother convergence to the solution.

In this work, we present the development of an efficient micro-kinetic model solver implementation in CatMAP. The program solves the mean-field micro-kinetic model to determine the steady-state coverage of all reaction intermediates. We reformulate the steady-state expressions to be solved as a function of the number of sites instead of the coverage. Through this transformation, we implicitly encode the constraint that the sum of the coverage of all species on a given surface must be 1 and that all coverages are non-negative. Effectively, we transform the problem from a constrained optimization/root finding problem to an unconstrained and generally more stably solvable optimization problem. The paper is structured as follows; in Section 2, we recap the procedure currently used to determine the steady-state coverage. In Section 3, we discuss our implementation of the numbers-based solver. Finally, in Section 4, we validate and test our implementation by computing the steady-state coverages of a few prototype reactions in heterogeneous catalysis.
2 CONVENTIONAL STEADY-STATE \( \theta \)-BASED SOLVER

Consider a reaction network having \( n \) species adsorbed on a surface with \( N \) sites. The coverage of each species is determined by the number of sites it occupies. For a species \( i \) occupying \( n_i \) sites, the coverage is \( \theta_i = n_i/N \). The sum over the coverages of all \( n \) species plus that of the vacant sites \((n_\star/N)\) is required to be exactly 1,

\[
\sum_{i=1}^{n} \theta_i + \theta_\star = 1, \quad 0 \leq \theta_i \leq 1, \quad 0 \leq \theta_\star \leq 1
\]  

(1)

i.e. adsorbates can occupy at most the same number of sites as is available on the surface, and no single species can occupy more sites than available (nor can it have a negative coverage, i.e., occupy a negative number of sites). The coverages of all \( n \) species are related to each other through the reactions which occur on the surface. Consider \( M \) reactions in total, each reaction \( m \) with a forward rate given by \( r_m \) and a reverse rate by \( r_{-m} \). The rate of change (with time \( t \)) of \( \theta_i \) is given by,

\[
\frac{\partial \theta_i}{\partial t} = \sum_{m=1}^{M} s_m r_m (\theta_1, \theta_2 \ldots \theta_i \ldots \theta_n, \theta_\star) - \sum_{m=1}^{M} s_{-m} r_{-m} (\theta_1, \theta_2 \ldots \theta_i \ldots \theta_n, \theta_\star)
\]  

(2)

where \( s_m \) and \( s_{-m} \) are the stoichiometric indices of the species \( i \) in a reaction \( m \). For simplicity, we refer to the combined function of the rate of formation of a species \( i \) as \( f_i \).

\[
\frac{\partial \theta_i}{\partial t} = f_i (\theta_1, \theta_2 \ldots \theta_i \ldots \theta_n, \theta_\star)
\]  

(3)

Under steady-state conditions, the coverages are unchanged with time,

\[
0 = f(\theta)
\]  

(4)

where \( f : \mathbb{R}^{n+1} \mapsto \mathbb{R}^M \) maps from the vector \( \theta \) of all coverages, including the empty site coverage \( \theta_\star \), to \( f_i \) (boldface symbols represent a vector of all \( i \)). Note that there are only \( n \) independent equations in (4), and not \( n+1 \), as the rate of change of \( \theta_\star \), \( f_\star \), is the negative of \( \sum_{i=1}^{n} f_i \) due to the constraint (1).

Equation 4 represents a steady-state micro-kinetic problem composed of a set of coupled, non-linear algebraic equations, where the goal is to find \( \theta \). Conventional solution procedures
solve Equation 4 for \( n \) species, while \( \theta_* \) is directly taken from Equation 1, i.e. by subtracting the coverage of all species from 1. This particular solution strategy requires a constrained root-finding algorithm to find \( \theta \), as the physical range of all \( \theta \) (including \( \theta_* \)) are constrained to \([0, 1]\).

In practice, \( \theta \) is obtained through a Newton root-finding approach based on variables described by arbitrary precision.\(^{16}\) Starting from an initial set of guesses at \( \theta_{\text{init},i} \), given by the Boltzmann coverage,

\[
\theta_{\text{init},i} = \frac{\exp \left( \frac{-G_i}{k_B T} \right)}{\sum_{i=1}^{n} \exp \left( \frac{-G_i}{k_B T} \right)},
\]

where \( G_i \) is the free energy of formation of species \( i \), \( k_B \) is Boltzmann’s constant and \( T \) is the temperature. Due to the exponential dependence of the coverages on the free energies \( G_i \), the coverages \( \theta_i \) can differ by many orders of magnitude, requiring the arbitrary precision arithmetic for the solution of Equation 4.

At each iteration with index \( a \), \( \theta_a \) is altered by \( \delta \theta \), following

\[
\mathbf{J}(\theta_a) \cdot \delta \theta = -\mathbf{f}(\theta_a)
\]

where the \( \mathbf{J}(\theta) \) is the Jacobian of \( \mathbf{f} \). This procedure is carried out until \( \mathbf{f}(\theta) = 0 \) or a stationary point is reached (i.e. the root-finding search fails).

The above procedure to determine \( \theta \) constitutes a constrained minimisation / root-finding problem due to the constraints in Equation 1. While it is intuitive to solve for \( \theta \) in \( \theta \)-space, the constraints in Equation 1, which in the original CatMAP implementation are imposed by rescaling the coverages after each Newton step, can lead to poor convergence in practice. We note in passing that there are other approaches to reach the steady state solution, such as solving the coupled differential equations \( \mathbf{f}(t) \) and extrapolating till \( t \to \infty \).\(^{11}\)

### 3 NUMBER OF SITES, \( x \)-BASED SOLVER

In this section, we develop an approach which circumvents the need to explicitly enforce the constraints in Equation 1. We do so by working in the space of the number of sites, \( n \),
instead of $\theta$. Specifically, we express $\theta$ as,

$$
\theta_i = \frac{x_i^2}{\sum_{j=1}^{n} x_j^2}
$$

(7)

where $x_i^2 = n_i$ plays the role of the number of sites and $\sum_{j=1}^{n} x_j^2 = N$ (total number of sites). The index $i$ is over both adsorbates and sites as opposed to the previous section, which is only over adsorbates. Written in this form, the steady state problem solves for $x$ (vector of all $x_i$) instead of $\theta$. There are no constraints on the $x_i$; Equation 7 ensures that all coverages $0 \leq \theta_i \leq 1$, and that the coverages sum to unity.

Note that by solving for all $x_i$, including the number of free sites $x_*$, we have an under-determined system of equations, i.e. the number of variables is more than the number of independent equations ($n + 1$ variables including $x_*$ and $n$ independent equation in $f$).

Effectively, this transformation allows us to solve a linear least-squares fitting problem for all $x$ instead of the conventional constrained root-finding problem presented in Section 2.

## 3.1 Modification of variables for calculations in $x$-space

We now describe the modifications to the conventional solution procedure (described in Section 2) that are required to solve the micro-kinetic model in $x$-space. We will continue to use the Newton-root procedure, analogous to Equation 6, just in $x$-space,

$$
J(x_a) \cdot \delta x = -f(x_a)
$$

(8)

where $a$ is the index of the iteration. The initial guess for $x$ is performed through the “Boltzmann-numbers”, analogous and in practise equivalent to the Boltzmann-coverage in Equation 5, without the normalisation,

$$
x_{i,\text{init}} = \exp\left(-\frac{G_i}{2k_BT}\right)
$$

(9)

Equation 7 states that $\theta$ is a function $g(x)$ of $x$ ($\theta = g(x)$), so it follows that the steady-state function in Equation 4 is $f(x) = f(g(x))$.

To convert $J(\theta)$ to $x$-space $J(x)$, we first define a matrix $M$,
\[
M = \begin{bmatrix}
\frac{\partial \theta_1}{\partial x_1} & \cdots & \frac{\partial \theta_1}{\partial x_i} & \cdots & \frac{\partial \theta_1}{\partial x_n} \\
\frac{\partial \theta_i}{\partial x_1} & \cdots & \frac{\partial \theta_i}{\partial x_i} & \cdots & \frac{\partial \theta_i}{\partial x_n} \\
\vdots & & \ddots & & \vdots \\
\frac{\partial \theta_n}{\partial x_1} & \cdots & \frac{\partial \theta_n}{\partial x_i} & \cdots & \frac{\partial \theta_n}{\partial x_n}
\end{bmatrix}
\]  
(10)

where, by chain rule of Equation 7, we get each element of \( M \) as,

\[
\frac{\partial \theta_i}{\partial x_k} = 2x_i \frac{2}{\sum_{j=1}^{n} x_j^2} \cdot \delta_{ik} - \frac{x_i^2}{\left(\sum_{j=1}^{n} x_j^2\right)^2} \cdot 2x_k
\]  
(11)

where \( \delta \) is the Kronecker delta function. Using \( M \), we construct an expression for \( J(x) \),

\[
J(x) = J(\theta) \cdot M
\]  
(12)

### 3.2 Modification of solution procedure

To obtain the step size \( \delta x \), we invert \( J(x_a) \) to obtain \( J(x_a)^{-1} \),

\[
\delta x = -J(x_a)^{-1} f(x_a)
\]  
(13)

As we have an under-determined system of equations, we use the Moore-Penrose pseudo-inverse to estimate \( J(x)^+ \)

\[
J(x)^+ = V \cdot S^+ \cdot U^T
\]  
(14)

In practice, we first perform a Singular Value Decomposition (SVD), as \( J(x) = U \cdot S \cdot V^T \).

We then determine the diagonal \( S^+ \), which is the inverse of the diagonal matrix \( S \) for non-zero diagonal elements, and zero elsewhere. This transformation gives the desired \( \delta x \),

\[
\delta x = -J(x_a)^+ f(x_a)
\]  
(15)

Finally, the new guesses, \( x_{a+1} \) are generated from the previous guess \( x_a \) by,
\[ x_{a+1} = x_a + l \cdot \delta(x) \]  

where \( l \) is a damping factor. \( l \) is reduced from 1 by a factor of 2 (i.e. 1, \( \frac{1}{2}, \frac{1}{4}, \ldots \)) until

\[ |f_{x_{a+1}}| \leq |f_{x_a}| \]  

4 TESTING SOLVER PERFORMANCE

Figure 1: Least squares errors computed through (left) the coverages solver and (right) the numbers solver for all iterations taken by the solver; both errors are generated for the CO-hydrogenation micro-kinetic model with a decimal precision of 300. Each curve represents a distinct set of formation energies \( G_i \) generated by varying the free energies of the descriptor (see text). The colours indicate if the solver converged to the threshold of \( 10^{-10} \) (in green) or failed to converge (in red). The x axis has been truncated to 20 iterations.

In this section, we test our implementation of the numbers solver on a few prototype reactions in heterogeneous catalysis. In particular, we compare the convergence behaviour
and the number of successful and failed solver attempts for four reactions from previous studies, ideal (interaction-less) CO oxidation,\textsuperscript{5,17} CO oxidation with adsorbate-adsorbate interaction parameters,\textsuperscript{5} thermal CO hydrogenation,\textsuperscript{4,18} and thermal methanol synthesis.\textsuperscript{18} We provide the relevant reaction details and input files to run the reaction network in Section S3 and defer mechanistic details to the cited publications.

As in the current CatMAP implementation, we independently vary the free energies of two chosen intermediates for all reactions and let the values of the reaction and activation free energies for all other intermediates change based on computed linear scaling relations.\textsuperscript{5} For example, in the case of CO-oxidation, the free energies of O* and CO* are varied between $-1 \text{ eV}$ to $3 \text{ eV}$ and between $-0.5 \text{ eV}$ to $4 \text{ eV}$ respectively, while the activation energies for O−O splitting and O−CO formation, which are elementary steps in the reaction network, are determined from linear scaling relations (see Section S1 for further details). As in the original CatMAP implementation, we vary the two descriptors on a uniform grid, generating different formation energies $G_i$ for all intermediates. We then solve the steady state micro-kinetic model to determine $\theta$ based on these $G_i$ values using both the coverage and numbers solver. Throughout the testing process, we use consistent solver parameters between the coverages and numbers solver, i.e., a tolerance of $10^{-10}$ and allowing for damping through Equation 16 up to 15 times; we leave the decimal precision as a free parameter and vary it during the testing procedure.

We present solver performance starting only from Boltzmann guesses. Using a 10x10 grid for all test cases, we expect that each point on the grid either succeeds or fails with no allowed option for restarting the solution procedure with an alternative initial guess. We note that the number of failed solutions for either solver can be reduced by attempting to re-use solutions from similar $G_i$ as initial guesses in an iterative procedure spanning the whole coverage maps as a function of the $G_i$.

As a metric to determine the progress towards the solution, we determine the error at each iteration as the least squares norm of the rate of change of $\theta_i$ (i.e. the steady state function),
\[
\text{Error} = \sum_i \left( \frac{\partial \theta_i}{\partial t} \right)^2 .
\] (18)

We choose this metric for the numbers solver as it allows us to vary the damping factor along the same norm that the Jacobian \( J(x) \) is computed through for the least-squares solver. We note that the original CatMAP code uses the maximum of all \( \partial \theta_i / \partial t \) to determine convergence, while our implementation uses the least-squares norm to determine the stepsize \( l \) in Equation 16.

Figure 1 shows the variation of the least-squares error for a given set of \( G_i \) for the methanol synthesis reaction. We consider the system to have converged when the corresponding error in coverage is less than the tolerance of \( 10^{-10} \) (points in green). Any sequence of iterations with an error larger than this tolerance is considered to have failed the convergence test (points in red). We find that the numbers solver (in Figure 1 b) converges exponentially for most points (green points decrease linearly on a log scale). At the same time, the coverage solver fails to converge and does not progress beyond 15 iterations due to the failure to meet the requirement of Equation 17. Similar plots for CO oxidation and CO hydrogenation are shown in Section S1.

Figure 2 shows the number of successful attempts by the solver to determine a solution for a given set of descriptor values for the four prototype reactions considered in this work. To test the stability of both solvers, we report the number of successful attempts against the decimal precision, which is varied from 200 to 550. The micro-kinetic models are of increasing complexity from left to right (a to d). For the simplest model, both solvers perform well, with the coverages solver (orange) solving 94/100 points while the numbers solver (blue) converges all 100 points. Including adsorbate-adsorbate interactions in the CO-oxidation model significantly lowers the number of successful solutions to 42/100 for both the coverage and the numbers solver.

The benefits of the numbers solver become evident with the increasing complexity of the micro-kinetic models. In the case of CO hydrogenation, the coverages solver converges 0/100 points, while the numbers solver still finds a solution for 28/100 points. For the methanol synthesis micro-kinetic model, the number of points converged depends on the chosen precision. The coverages solver appears to be unstable upon increasing the precision...
to greater than 300, with the number of solved points decreasing from 28/100 to 0/100. In contrast, the numbers solver presents solutions for 27 to 43/100 points for all precision values considered.

![Figure 2: Number of successfully solved points for the numbers (blue) and coverages (orange) solver; the maximum possible successfully solved points is 100.](image)

Based on our tests of four prototype reactions, we find that the numbers solver outperforms the coverage solver and allows for a stable approach to convergence. We envision its use in complex reaction networks where convergence is hard to achieve with the coverage-based solver.

5 **CONCLUSION**

In this work, we developed a method which computes the steady state coverage in microkinetic models. Our solver operates in the space of *number of sites* as opposed to that of *coverage*. Working in the space of *number of sites* absolves us from the need to explicitly constrain the coverages of all species on a surface to sum to 1 and the individual coverages to have physical values between 0 and 1, hence improving convergence. This transformation of variables converts the problem from a constrained minimisation / root-finding problem to a least-squares minimisation problem. We discuss the implementation of the solver routine into the open source kinetic modelling package CatMAP and show the significantly improved convergence of our numbers solver versus the coverage solver using a few representative reaction networks in heterogeneous catalysis.
ACKNOWLEDGMENTS

Support to the SUNCAT Center for Interface Science and Catalysis by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program is gratefully acknowledged. The research leading to these results has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 851441, SELECTCO2, and from The VILLUM Centre for the Science of Sustainable Fuels and Chemicals (9455) from VILLUM FONDEN. We thank Dr. Ahmed Elnabawy for valuable discussions about CO-hydrogenation kinetic modelling and Dr. Georg Kastlunger, Dr. Ang Cao and Sihang Liu for testing the code and providing valuable feedback.

DATA AVAILABILITY

All kinetic models in this work have been evaluated using the open source kinetic modelling package CatMAP. Source code for our implementation as well as log files for all tests are available at: https://github.com/sudarshanv01/catmap-mirror.

References

1. Schlexer Lamoureux, P.; Singh, A. R.; Chan, K. ACS Catalysis 2019,


