UNCERTAINTY QUANTIFICATION IN MULTISCALE STOCHASTIC MODELS OF CATALYTIC REACTIONS

by

Marcel Nunez

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering

Winter 2018

UNCERTAINTY QUANTIFICATION IN MULTISCALE STOCHASTIC MODELS OF CATALYTIC REACTIONS

by

Marcel Nunez

Approved: _____

Eric M Furst, Ph.D.
Chair of the Department of Chemical and Biomolecular Engineering

Approved: _____

Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Dionisios G. Vlachos, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Antony N. Beris, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Yushan Yan, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Markos A. Katsoulakis, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

This material is based upon the work supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, and Applied Mathematics program under Award No. DE-SC0010549 and the Defense Advanced Research Projects Agency (DARPA) under Award No. W911NF-15-2-0122.

In addition, I would like to acknowledge several of the colleagues who assisted my completion of this thesis. Early in my graduate career, Dr. Wei Guo, Dr. Matthew Christiansen, Vassili Vorotnikov, and Dr. Jonathan Sutton all served as patient mentors. Dr. Jeff Frey provided essential help enabling my computer codes to run on the university clusters. My classmates within the Vlachos group, Geun Ho Gu, Dr. Ryan Patet, and Alexander Mironenko, have been motivating companions. It has been a pleasure to share this research experience with them.

On the personal side, I thank my parents, Susan and Cesar Nunez, for the innumerable sacrifices they made for my education. They instilled in me the faith necessary to persevere through the inherent uncertainties of graduate school and scientific research. Their work ethic has served a model for me in graduate school and will continue to do so throughout the entirety of my professional career.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Multiscale modeling, a key tool in probing the fundamentals of catalytic reactions, has seen increased usage enabled by advances in computational hardware. Within the multiscale modeling paradigm, kinetic Monte Carlo (KMC) is employed to simulate chemical reaction networks, as mean-field models often fail to provide a meaningful description of the complex phenomena involved. Due to KMC's high computational cost and stochastic noise, quantifying uncertainty for the purposes of refining the model and assessing predictive reliability is difficult. Uncertainty arises from errors in input parameters (parametric uncertainty) and assumptions made about the physical system (model form uncertainty).

In this thesis, we develop tools to quantify errors from each of the aforementioned sources and make recommendations for model refinement. We address parametric uncertainty by developing efficient sensitivity analysis techniques, which identify the most influential parameters. Likelihood ratio sensitivity analysis (LRSA) computes all sensitivities without the need for additional runs, as required by finite difference methods, but encounters tremendous variance in systems with disparate time scales. To overcome this limitation of LRSA, we derive mathematical theory that enables its use in well-mixed multiscale KMC and implement the method in original software. The new multiscale technique accurately computes sensitivities in a model system for which the traditional LRSA performs poorly. To address spatial KMC, we develop acceleration techniques and statistical criteria that ensure sufficient sampling for LRSA. As a result, LRSA can be applied to real chemistry. We apply our methodology to the water-gas shift reaction on Pt, an important component of hydrogen production from biomass.

We address model form uncertainty by revisiting two common assumptions: the structure of the catalyst surface is uniform and the identity of the active site is known.

A framework for optimizing catalyst structure based on local descriptors is developed, allowing for atom-by-atom design of defected surfaces and consequent improvements in activity. In order to restrict our search to physically relevant structures, surface energy is also computed. Activity is maximized and surface energy is minimized simultaneously using multi-objective simulated annealing. A set of Pareto optimal structures is found, offering targets for synthesis. We apply our approach to oxygen reduction on Pt, the key reaction in automotive fuel cells. Our approach resolves discrepancy between experiment and theory regarding the extent to which defects can improve activity. We extend the approach to chemistries involving coupled active sites, for which KMC simulation is needed. KMC simulation data from many different structures is used to train a neural network for use as a surrogate model in the optimization. The neural network is updated as the optimization progresses in an online machine learning approach. In doing so, geometric effects such as diffusion limitations and bifunctional site coupling are accurately captured within the structure optimization. The impact on the optimal structure is analyzed, yielding new insights into catalyst structure/activity relationships.

Chapter 1

INTRODUCTION

In this chapter, we motivate the thesis work and provide relevant background material. Section 1.1 discusses the need for developing new catalytic materials based on a fundamental understanding of catalytic phenomena. We contend that computational multiscale modeling is crucial to this objective and describe the approach. Section 1.2 provides an overview of the kinetic Monte Carlo method, a key component of the multiscale modeling methodology. Section 1.3 discusses the role of uncertainty quantification in multiscale modeling and the aspects that are lacking for kinetic Monte Carlo. Section 1.4 defines the objectives and scope of the thesis.

1.1 Multiscale Modeling of Catalytic Reactions

Catalyzed chemical transformations have made possible many of the modern comforts we enjoy today. The inventions of nylon clothing, petrochemical derived plastics, synthetic rubber, transportation fuels, and many other everyday commodities all featured technological advances in catalytic processes[13]. Nevertheless, rapidly increasing global demand for energy[2] and materials will strain existing processes in the near future. Meeting this demand in an economically and environmentally sustainable manner will require further advances in catalytic technology[1]. Three areas ripe for improvement are as follows. First, the production of ammonia for fertilizer via the Haber-Bosch process consumes 2% of our global energy[54]. A catalyst requiring lower temperatures and pressures would drastically reduce the energy demand of this process. Second, many of the environmental and geopolitical disadvantages of petroleum could be avoided by switching to natural gas[194] and biomass[90, 24] as a feedstock for fuels and chemicals. However, new catalytic processes will be required, as the chemical

compositions of these alternative feedstocks differ from petroleum. Third, a cheap, stable catalyst able to efficiently reduce oxygen to water in the cathode of automotive fuel cells[91, 92] would enable cars to have quick refueling, long driving range, and no tailpipe emissions. In summary, the array of chemistries and applications involved is so extensive that no single technological breakthrough will address them all. Rather, an improved fundamental understanding of catalytic reactions is necessary for the design of new catalysts.

Historically, catalysts have been developed in a largely empirical fashion[124, 232, 12]. Future advances will require a more fundamental and robust approach. Analytical techniques such as spectroscopy[39], microscopy, gas chromatography, BET surface area measurement[25], etc. are now staples of modern catalysis[43] and have helped gain useful insights. However, experimental techniques have limitations in their spatial and temporal resolutions. Furthermore, they cannot always study the catalyst under reaction conditions[84]. A promising complementary technique is first principles (computational) modeling[150, 27, 107, 176], which describes a catalytic system based on fundamental, physics-based equations. It can, for example, test the feasibility of reaction mechanisms proposed by experimental studies based on the electronic structures of the molecules involved. The computational approach circumvents the issue of spatial and temporal resolutions by using the multiscale modeling approach discussed below. In addition, computational modeling benefits from the consistent increases in computer power following Moore's law[127], resulting in ever increasing system sizes that can be studied. Computations typically allow for studies at much lower financial cost than what would be required for comparable experiments.

The multiscale modeling methodology[183, 224] leverages computational hardware in an efficient manner towards the inherently multiscale problems in catalysis. The brilliance of the multiscale modeling methodology is that it employs different tools to study the different length and time scales. Information is exchanged between the disparate scales to couple them. In doing so, macroscopic observables are predicted using fundamental information about molecular properties and behaviors. A schematic

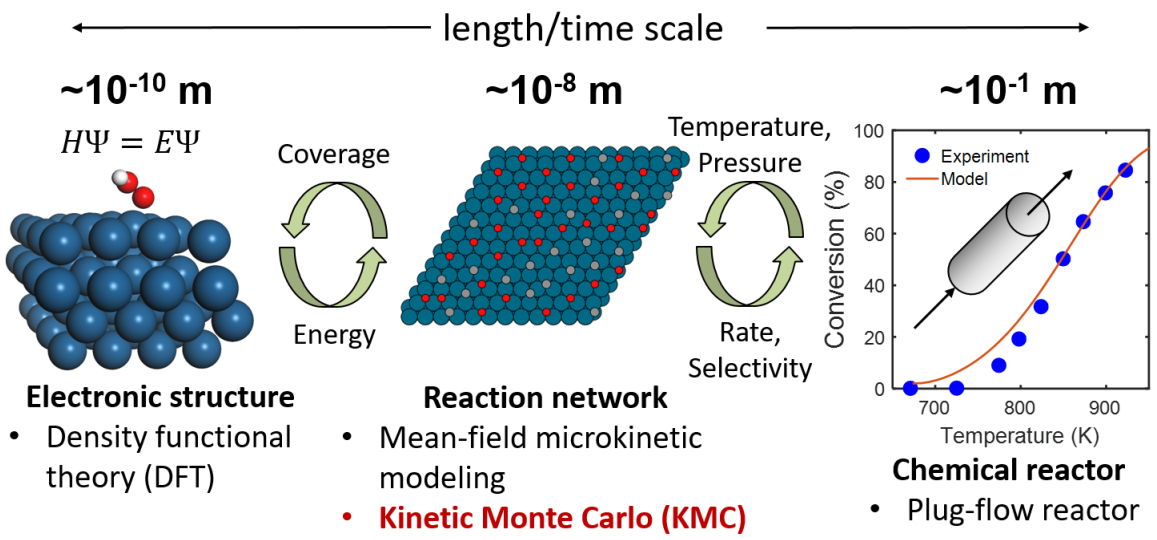


Figure 1.1: Overview of the multiscale modeling paradigm. Information is passed up and down scales to link macroscopic observables to fundamental molecular properties.

of multiscale modeling is shown in Figure 1.1.

At the quantum scale (angstroms), the electronic structure of atoms and molecules is probed with ab initio methods such as density functional theory (DFT)[110]. DFT solves the Kohn-Sham equations, which are derived using approximations to the Schrödinger equation. The Hohenberg-Kohn theorems state that all properties of the system, such as energy, can be expressed as a functional of the electron density. In essence, the Kohn-Sham equations reduce the complexity of the many-electron problem by solving the single electron problem in an effective potential determined by the electron density. The electron density and effective potential must be computed by solving a self consistent set of equations. The origin of this method from a fundamental equation describing the behavior of the electrons qualifies it as a first principles method. DFT describes the electronic behavior to the extent that it obtains numerical accuracy in solving the Kohn-Sham equations and that the exchange-correlation functional captures interactions between the electrons. Within the DFT framework, there is a hierarchy (termed after the Biblical "Jacob's ladder") of exchange-correlation functionals with tradeoffs between accuracy and cost[197].

This thesis focuses specifically on extended transition metal surfaces, for which the generalized gradient approximation (GGA) functionals offer a suitable tradeoff[161]. There is extensive ongoing work to improve the accuracy of DFT. A few approaches are constraining the functional to theoretical limits[209], benchmarking to experimental data[231], and using machine learning[23].

Given a molecular configuration, DFT software (e.g. VASP) optimizes the coordinates of the nuclei of the atoms to minimize the total electronic energy of the system, allowing the calculation of thermodynamic parameters such as binding energies. Transition states of chemical reactions can be obtained using several methods such as the climbing image nudged elastic band method[83]. Vibrational frequencies contribute to the entropy of the system at finite temperatures and are found by perturbing atomic coordinates from the optimal configuration.

The CPU requirements of DFT depend on many numerical parameters and system specifications, but a typical calculation with a few dozen atoms takes several hours. For large chemical reaction networks where many thermodynamic and kinetic parameters must be computed, this cost becomes prohibitive. As an alternative, DFT data can be substituted by estimates obtained from semi-empirical methods regressed to DFT data. Group additivity[182, 69] and linear scaling relations[3] estimate thermodynamic properties. Brønsted-Evans-Polanyi (BEP) relations[16] and transition state scaling relations[229] estimate reaction barriers.

The macroscopic scale of multiscale modeling uses chemical reactor models. The field of chemical reactor engineering has been a core subject of the chemical engineering discipline since it was established as an official academic subject by the Massachusetts Institute of Technology in 1888[21]. Reactor models describe the macroscopic behavior of a reactant/product mixture as affected by the catalyzed reactions (uncatalyzed reactions may also occur). The phase of the feed may be gaseous or aqueous. The reactor model used should match the experimental system as closely as possible. In the simplest case, a 1-dimensional plug flow reactor model suffices. When necessary, more complex computational fluid dynamics models[80] are used and are coupled with

the surface chemistry. The system in inherently multiphysics, as heat transfer, mass transfer, and chemical reactions are all present. Solving the coupled problem allows for the computation of conversions, selectivities, reaction orders, and apparent activation energies under realistic conditions. These quantities are suitable for direct comparison to experimental results[17].

The quantum and macroscopic scales must be coupled for electronic information to affect predictions of macroscopic observables. The reaction network scale (center of Figure 1.1) accomplishes this task. Here, electronic information is used, along with statistical mechanics, to parameterize the rate constants of elementary chemical reactions. Information is fed back down to the quantum level when the reaction network simulation informs the DFT of what are the relevant molecular configurations to compute. Reaction rates computed by the chemical reaction network simulation are fed up to the macroscopic reactor model and are incorporated into the multiphysics simulation. Macroscopic reactor conditions such as temperature and partial pressures are fed back down, as they affect the rates of elementary reactions. Networks of elementary chemical reactions are simulated using either mean field microkinetic modeling (MKM) or kinetic Monte Carlo (KMC). This thesis develops methods for kinetic Monte Carlo, as discussed in the next section.

1.2 Kinetic Monte Carlo (KMC)

Mean-field modeling is used extensively to model surface catalyzed reactions, but it relies on approximations that frequently break down for systems of interest[183]. It assumes that the spatial distribution of adsorbates on the surface is uniform. However, lateral interactions, spatial correlations, and finite diffusion rates affect the availability of neighboring reactant pairs in bimolecular reactions[213, 94, 5]. Furthermore, the continuum approximation made by mean-field models is only valid for large numbers of particles, which is not the case on small nanoclusters[202]. KMC avoids these issues by tracking each species on the surface individually, thus obtaining atomistic resolution and accurate statistical averaging. For these reasons, KMC has

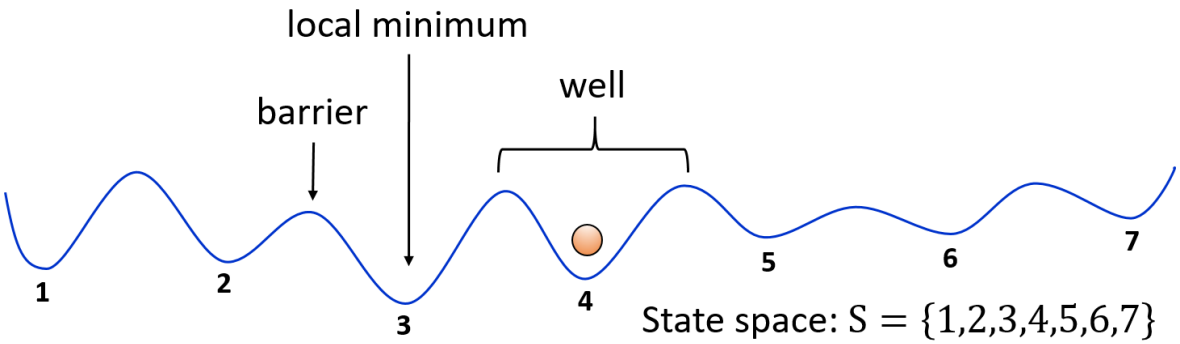


Figure 1.2: Diagram of a 1-dimensional potential energy surface. The currently occupied state is indicated by the orange ball. The state space (S) enumerates all local minima on the potential energy surface.

been applied to a wide range of surface chemistries[175, 206, 200]. In this section, we provide a brief overview of the KMC method as applied to surface catalyzed chemical reaction networks.

1.2.1 Theory

For a given chemical system where the positions of all atoms are known, the (high dimensional) potential energy surface (PES) describes the potential energy of the system as a function of the x, y, and z coordinates of each atom. The PES can be computed through ab initio methods such as DFT or an appropriate interatomic potential. A stable chemical configuration is a local minimum on the PES. Energy wells describe the area near a local minimum and are separated by barriers. Thermal fluctuations cause the system to move around the PES, sometimes crossing a barrier to an adjacent energy well (neighborhood of a local minimum). The crossing of a barrier defines a chemical reaction (e.g. bond breaking/formation, diffusion, adsorption/desorption). Transitions between wells of the PES occur much less frequently than fluctuations within the well. In contrast to molecular dynamics (MD) simulation, KMC coarse-grains the potential energy surface by tracking only the local minima. Figure 1.2 shows a cartoon of a 1-dimensional PES where the local minima are labeled as discrete state.

Crossing between local minima requires traversing a transition state (TS), i.e. a saddle point on the PES. A transition state is a local minimum in all but one dimension (after a suitable change of coordinates), in which it is a local maximum. Transition states are not explicitly included in KMC simulation, but are used to compute reaction propensities, or rates of transition between discrete states. Mathematically, the propensity is the inverse of the expected waiting time until the reaction spontaneously occurs. The propensity of the reaction transitioning between an initial state (IS) and final state (FS) is calculated using statistical mechanics and transition state theory as

$$a_{\mathrm{IS}\to\mathrm{FS}} = \frac{k_B T}{h}\frac{Q^{\mathrm{TS}}}{Q^{\mathrm{IS}}}\exp\left(\frac{E_{\mathrm{TS}} - E_{\mathrm{IS}}}{k_B T}\right). \tag{1.1}$$

Q denotes the partition function taking into account translations, vibrations, and rotations. E denotes the electronic energy. k_B and T denote Boltzmann's constant and temperature respectively. Further details are available in a review by Stamatakis and Vlachos[206].

Chemical reaction networks have many states arising from the combinatorial arrangement of molecules in a system. The system occupies one state at a time. Let $p_i(t)$ be the probability that the system is in state i at time t. The time evolution of the system is governed by a continuous time Markov chain (CTMC) called the Chemical Master Equation (CME)[64].

$$\frac{dp_i(t)}{dt} = \sum_{j\neq i}\left[a_{j\to i}p_j(t) - a_{i\to j}p_i(t)\right] \tag{1.2}$$

Equation 1.2 defines a linear set of ordinary differential equations (ODE). In principle, it can be solved analytically (see Section B.1) using a matrix exponential or numerically using linear multistep methods. However, for most practical problems, the cardinality of the state space (n_{states}) is too large for either of these approaches to be feasible. In fact, even enumerating the entire state space is often impractical. For the relatively simple case of adsorption of a single species (e.g. CO) on the top sites of a $p(10\times 10)$ Pt(111) surface, each site can either be occupied or vacant. Already there are $2^{10\times 10} = 2^{100}$ states of the lattice. Solving such a large number of equations would

be infeasible even with modern computers. Rather than fully solving Equation 1.2 for every state, we statistically sample the probability distribution function (PDF) using kinetic Monte Carlo simulation, discussed next.

1.2.2 Simulation

In KMC simulation, a trajectory is a randomly generated time series of states occupied by the system, with probability governed by the CME. The transition between states at a given time defines a KMC step, or reaction event. At any given step, the simulation must determine which state to visit next and at what time, according to an algorithm. For simplicity, we discuss the Gillespie algorithm[61]. Other algorithms achieve a mathematically equivalent PDF, but have different tradeoffs in terms of memory and CPU resources, as discussed in by Chatterjee and Vlachos[41]. All possible reaction events (i.e. events for which the propensity is nonzero) are enumerated in a reaction event catalogue and their propensities are calculated. Two random numbers are generated per step. The first random number is used to choose the time step (Δt) according to the exponential distribution as follows.

$$E[\Delta t] = \left(\sum_{k \neq i} a_{i \to k} \right)^{-1} \tag{1.3}$$

$$P(\Delta t \leq t') = 1 - \exp\left(-\frac{t'}{E[\Delta t]} \right) \tag{1.4}$$

$E[\Delta t]$ denotes the expectation value of the time step. $P(\Delta t \leq t')$ denotes the probability that the time step is less than t', thus defining a cumulative probability distribution. The second random number chooses the reaction to fire. $P(i \to j)$ denotes the probability that j is the next state visited after i.

$$P(i \to j) = \frac{a_{i \to j}}{\sum_{k \neq i} a_{i \to k}} \tag{1.5}$$

After a reaction fires, the reaction event catalogue is updated. In practice, not all propensities need to be recalculated. There may be processes which transition between different states, but represent the same physical process and therefore have the same

propensity. For example, adsorption of a molecule onto a lattice site is unaffected by reaction events occurring at a distant lattice site. For any reaction events which can no longer occur, they are removed from the reaction event catalogue. Then, new reactions that can now occur but were not available before are identified. The propensities of the newly identified reactions are computed and are added to the reaction event catalogue. A pseudocode of the KMC algorithm is as follows.

1. Enumerate all possible reactions in a reaction event catalogue.
2. Compute the propensity of all reactions in the reaction event catalogue.
3. Choose a time step Δt according to Equations 1.3 and 1.4.
4. Choose a reaction according to Equation 1.5.
5. Fire the reaction, updating the system state.
6. Increment the time clock by Δt.
7. Remove reaction events that can no longer occur from the reaction event catalogue.
8. Identify new reaction events that can occur and compute their propensities. Add them to the reaction event catalogue.
9. Iterate until the termination time (t_f) has been surpassed.

1.2.3 Lattice KMC

Well-mixed KMC uses mass action kinetics in a similar way to mean-field modeling, with the exception that a discrete number of particles is used. This version of stochastic simulation is popular in the biology community[11]. In heterogeneous catalysis, reactions occur on surfaces, where the spatial distribution of adsorbates is important. Therefore, lattice KMC[20] is used extensively. The state of the system is defined by the occupancies of adsorbates on discrete lattice sites. Due to the crystallinity of metals and oxides, lattice sites can be straightforwardly defined. By distinguishing between different site types (e.g. top, hollow, step, etc.), KMC is able to describe heterogeneous surfaces while preserving the full spatial resolution.

Figure 1.3 shows how for a Pt(111) surface, the continuous x, y, and z coordinates of all the atoms are coarse grained to discrete, well defined lattice positions. This makes it possible to identify patterns on the surface with graph theory. The energies of different configurations are computed through a cluster expansion Hamiltonian[239, 190, 234]. The lattice state is defined by a vector of lattice occupancies. In our example, we order the sites in a dictionary fashion with the 9 top sites preceding the 9 fcc hollow sites. We indicate the occupancies as being a vacancy (0), H_2O^* (1), OH^* (2), or H^* (3). Following our definition, the change in lattice occupancies due to the reaction in Figure 1.3 is as follows.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The ability of lattice KMC to change the types, quantities, and spatial arrangement of surface sites makes it an integral tool for studying structure sensitivity. Terraces, steps, 3D nanoparticles, and defects can all be modeled. Throughout this thesis, we use the Zacros[204, 148] graph-theoretical code for lattice KMC, which has been employed in several such studies[201, 202, 130, 73, 164, 149, 75, 203, 165].

1.3 Uncertainty Quantification (UQ)

Because multiscale modeling is such an intricate, multifaceted paradigm, naïve utilization of the most accurate methods on every scale would be computationally prohibitive. To manage this complexity in practice, a hierarchical refinement approach[183] is used. Initial attempts to model a system begin with a crude model, employing many approximations. Then, uncertainty quantification (UQ) identifies the most influential aspects of the model. Once identified, the influential aspects of the model are refined by using higher levels of theory or by gathering additional data, either through additional calculations or experiments. In doing so, the model is iteratively improved until it is a satisfactory representation of the true system (i.e. it achieves reliable predictions)

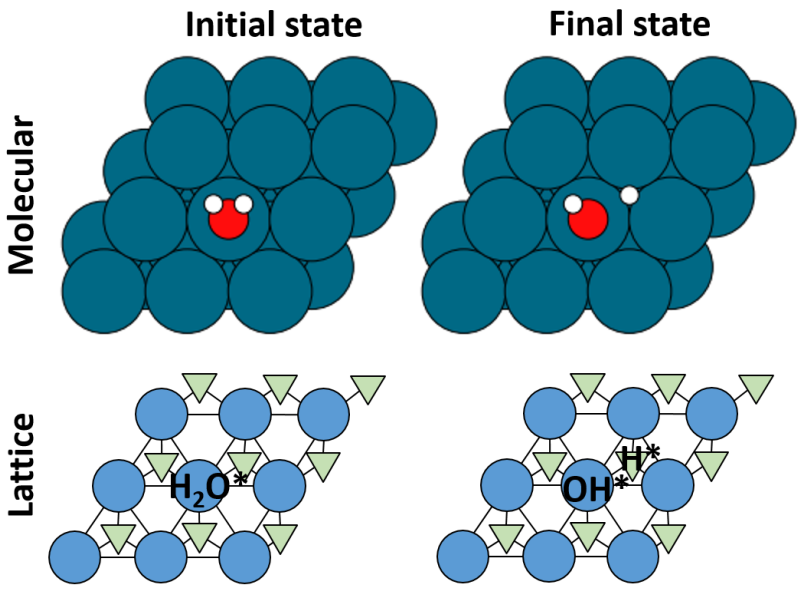


Figure 1.3: Molecular (top row) versus lattice (bottom row) representations of a catalyst surface. In the molecular picture, all atoms have continuous x, y, and z coordinates. Colors indicate Pt(blue), O(red), and H(white) atoms. The lattice is a discretization of the full molecular picture. The reaction shown is H_2O^* dissociation into OH* and H* on a Pt(111) surface. In the KMC lattice, blue circles are top sites and green triangles are fcc hollow sites.

is achieved. This approach allows for an efficient allocation of effort in constructing a multiscale model.

UQ is a crucial element in the process of hierarchical refinement. There have been many developments with applications to catalysis. The errors of DFT have been explored extensively[120, 119]. The Bayesian error estimation functional (BEEF) has built-in error estimates and currently sees wide use[134, 231, 44, 88, 220]. Sutton et al. quantified the effects of errors in semi-empirical methods on activity/selectivity maps[211, 212]. Likewise, the effect of lateral interactions on materials predictions has been investigated[219]. Many applications of UQ to reaction networks aim to reduce the reaction mechanism with minimal impact on predictions[163, 222, 193]. On the reactor scale, UQ has been successfully applied to fluid dynamics[59].

For the purposes of this thesis, the UQ task is divided into two pieces.

1. Parametric UQ: Quantifying how errors in the input parameters of a model propagate to the model predictions.
2. Model form UQ: Understanding how assumptions made about the physical system being modeled affect the results.

Parametric UQ is performed using sensitivity analysis (SA). Given a model f with input parameters $\theta_1, \theta_2, ...$ and observables $y_1, y_2, ...$, sensitivity analysis quantifies the effect of the input parameters on the observables. SA can be local or global. Local sensitivities are computed as partial derivatives. The sensitivity ($S_{i,j}$) of observable i to parameter j is given by

$$S_{i,j} = \left(\frac{\partial y_i}{\partial \theta_j}\right)_{\theta_{k\neq j}}. \tag{1.6}$$

Global sensitivity analysis extends local sensitivity analysis by fitting a response surface to f, which accounts for nonlinearities in parameter dependencies and interactions between parameters. Rather than using fixed values for the parameters, global sensitivity analysis chooses their values from a probability distribution. We focus on local sensitivity analysis in this thesis.

SA provides physical insights by identifying the rate-determining steps of a reaction network. In addition, it also allows for the identification of sensitive parameters that may be refined by fitting experimental data or through more precise methods. For example, sensitive parameters estimated using semi-empirical methods can be replaced with those from more accurate DFT calculations. SA has also been used to identify which parts of the reaction network might be missing[210]. Given the complexity of reaction network models, the numerical computation of Equation 1.6 is often not straightforward and requires sophisticated mathematical approaches. While this approach has been successfully demonstrated for mean field models, the lack of efficient SA tools has hindered the extension of hierarchical parameterization to KMC models. A more extensive review of SA methods and their limitations for KMC is given in Sections 2.2 and 3.2.

Model form uncertainty is mathematically less well-defined than parametric UQ and often more challenging. Even if accurate parameters are used, a model will be incorrect if it does not model the correct physical process. There are many physical processes that occur in catalytic systems, but are often excluded from models. Mass transfer limitations in the reactor, solvent effects, and catalyst reconstruction are a few examples. In this thesis, we address uncertainty arising from incomplete knowledge of the catalyst structure and the identity of the active site. Pidko[166] provides excellent discussion of the tradeoff between model simplicity and accuracy regarding the active site.

1.4 Dissertation Scope and Structure

The objective of this thesis is to enable uncertainty quantification for KMC. We develop parametric UQ methods in Chapters 2 and 3. Chapters 4 and 5 explore model form uncertainty by assessing assumptions made about the catalyst structure and the identity of the active site.

In Chapter 2 we derive and implement a two time scale approach to performing

sensitivity analysis in stiff systems. The previously developed likelihood ratio sensitivity analysis method is shown to encounter insurmountable statistical noise in single time scale systems with disparate reaction time scales. To resolve the issue, we explicitly identify fast and slow reactions in order to partition the simulation into fast and slow scales. Derivatives are computed accurately in each scale and combined. As a result, the two time scale method accurately predicts sensitivities for a stiff reaction network without the need to perturb reaction parameters and repeat simulations. The method is demonstrated in an analytically solvable linear system.

Chapter 3 takes an alternative approach towards sensitivity analysis. Whereas the method developed in Chapter 2 is only applicable to well-mixed KMC, practical systems of interest are on-lattice. We address the challenges of efficiently conducting KMC simulations and performing accurate sensitivity analysis in systems with unknown time scales by employing two acceleration techniques: rate constant rescaling and parallel processing. We develop statistical criteria that ensure sufficient sampling of non-equilibrium steady state conditions. Our approach provides the twofold benefit of accelerating the simulation itself and enabling likelihood ratio sensitivity analysis, which provides further speedup relative to finite difference sensitivity analysis. As a result, the likelihood ratio method can be applied to real chemistry. We apply our methodology to the water-gas shift reaction on Pt(111), a key reaction in hydrogen production from biomass.

In Chapter 4 we quantify uncertainty in catalyst activity due to unknown catalyst structure by solving an inverse problem. Given a structure sensitive chemistry, we optimize the structure within constraints imposed by the thermodynamic stability of the surface. We identify active sites using coordination-dependent binding energies, as well as the density and spatial arrangement of such sites. Simultaneously, surface energy is minimized in a multiobjective optimization approach. We apply this methodology to the oxygen reduction reaction on defected Pt(111), Pt(100), Au(111), and Au(100) surfaces.

In Chapter 5 we extend the structure optimization approach of Chapter 4 to

chemistries where KMC is needed to accurately compute the reaction rate for a given catalyst structure. KMC accounts not only the presence of active sites, but also geometric effects such as diffusion limitations and bifunctional site coupling. However, running a KMC simulation at each step of an optimization is infeasible. To reduce the computational burden to an acceptable level, we use a neural network surrogate model trained on KMC data for the optimization. The neural network captures the complex geometric dependencies of the KMC simulation at comparably negligible computational cost. We apply our method is to a model chemistry analogous to the ammonia decomposition reaction on NiPt.

Chapter 6 summarizes the findings of this thesis and proposes future directions.

Chapter 2

EFFICIENT SENSITIVITY ANALYSIS IN MULTISCALE KINETIC MONTE CARLO

2.1 Abstract

Kinetic Monte Carlo simulation is an integral tool in the study of complex physical phenomena present in applications ranging from heterogeneous catalysis to biological systems to crystal growth and atmospheric sciences. Sensitivity analysis is useful for identifying important parameters and rate-determining steps but the finite-difference application of sensitivity analysis is computationally demanding. Techniques based on the likelihood ratio method reduce the computational cost of sensitivity analysis by obtaining all gradient information in a single run. However, we show that disparity in time scales of microscopic events, which is ubiquitous in real systems, introduces drastic statistical noise into derivative estimates for parameters affecting the fast events. In this work, the likelihood ratio sensitivity analysis is extended to singularly perturbed systems by invoking partial equilibration for fast reactions, that is, by working on the fast and slow manifolds of the chemistry. Derivatives on each time scale are computed independently and combined to the desired sensitivity coefficients to considerably reduce the noise in derivative estimates for stiff systems. The approach is demonstrated in an analytically solvable linear system.

2.2 Introduction

Kinetic Monte Carlo (KMC) simulation has been successfully applied in several complex systems, ranging from crystal growth to surface reactions to biological networks to atmospheric sciences and to spatially homogeneous (known also as the stochastic simulation algorithm (SSA) or Gillespie algorithm[63]) as well as spatially

distributed systems[20]. Various KMC algorithms and accelerated methods were recently discussed in Chatterjee and Vlachos[41]. With advances in ab initio methods, a burst of scientific activity is seen in catalysis in what is termed as ab initio KMC simulations with examples including the study of ethylene conversion to ethylidyne on Pd and Pt[4], the contribution of surface defects to methanol decomposition on Pd[123], structure sensitivity of the water gas-shift reaction on Pt[201], oxygen coverage effects on NO oxidation on Pt[234], and the influence on Ni island structure on Pt for the decomposition of ammonia[73] as reviewed in Stamatakis and Vlachos[206]. Similar activity is seen on biological networks[174].

Error analysis is a vital part of stochastic modeling, and has been addressed for certain approximations such as coarse-graining[102, 9, 100]. In particular, parametric sensitivity analysis (SA) provides physical insights by identifying the rate-determining steps of a reaction network that can be interrogated with experimental studies or guide development of better catalysts[168]. In addition, it also allows for the identification of sensitive parameters that may be refined by fitting experimental data or through more precise methods. For example, sensitive parameters estimated using semi-empirical methods can be replaced with those from more accurate density functional theory (DFT) calculations. In this manner, SA allows for appropriate allocation of computational effort within the hierarchical refinement approach to multiscale modeling[183, 224]. While this approach has been successfully demonstrated for mean field models, the lack of efficient SA tools has hindered the extension of hierarchical parameterization of KMC models.

Various techniques have recently been developed for SA of stochastic kinetic systems, such as polynomial chaos expansion (PCE)[103, 104, 105] and Fisher information matrix (FIM)[158, 159, 71]. The traditional finite differencing SA approach is simple to implement, as it does not necessitate any changes to the underlying simulation algorithm[169]. Its variations such as the common random number (CRN)[66, 172], common reaction path (CRP)[172], and coupled finite difference (CFD)[6, 7] methods

reduce the noise in stochastic differences[199]. However, since finite differencing calculates derivatives through multiple runs in which parameters are perturbed individually, the computational burden scales linearly with the number or parameters. Increasing the order of approximation of the derivative estimates requires additional runs for each parameter. Furthermore, finite differencing introduces bias due to the perturbation size, which must be sufficiently large to obtain statistically significant differences but small enough to minimize local truncation error[133]. Therefore, the computational cost for expensive KMC simulations with many parameters is prohibitively large.

The likelihood ratio (LRSA) method[66, 67, 147] based on the Girsanov measure transformation[167] has recently been introduced for KMC simulations[133] because it is independent of the number of parameters and is the method of choice for short, transient calculations. However, statistics are taken over many replicate trajectories, typically thousands. Furthermore, statistical noise accumulates over longer simulation runs making it unusable for steady-state simulations that are ubiquitous in kinetic studies. Recently, the likelihood ratio has been adapted to steady-state calculations, alleviating both issues[230].

Real systems often exhibit separation of time scales which plagues KMC implementation[41, 202] and introduces noise for derivatives computed with the likelihood ratio method, as will be shown in Section 2.5.3. Several hybrid stochastic methods have been developed to accelerate KMC simulations[35, 33, 170, 68, 186, 81, 62, 185]. Among these, singular perturbation identifies the time scales on which each reaction occurs and properly slaves the rates (propensities) of fast reactions on the manifold of the slow reactions[52] Gupta and Khammash show that sensitivities obtained from a model in which the time scales are separated can be used to approximate the sensitivities of an exact, stiff model[76]. In this paper, we combine the singular perturbation methodology with the LRSA method. We show that this allows us to overcome the limitations of the standard LRSA methodology and obtain accurate sensitivity coefficients in stiff systems.

2.3 Likelihood ratio sensitivity analysis (LRSA)

In this section, we review the details of the LRSA relevant to our work. Given an initial state for a KMC simulation and a stopping time t, it is of interest to compute the expectation of a system property $f(t)$ at time t. f can be any property of the system state, such as a species population. Computation of $\langle f(t) \rangle$ requires averaging over multiple simulations to sample Ω, the space of all possible trajectories between times 0 and t which start at the given initial state. Glynn[67] expresses the expectation as

$$\langle f(t) \rangle = \int_{\Omega} P_{\omega}(t) f_{\omega}(t) d\Omega \tag{2.1}$$

where $P_{\omega}(t)$ is a probability measure on the space Ω and $\omega \in \Omega$ is a trajectory.

Now suppose we want to know the sensitivity of $\langle f(t) \rangle$ to parameter θ. Typical examples for θ include a rate constant or an activation energy. From Equation 2.1, we compute the derivative of the expectation as

$$\frac{\partial \langle f(t) \rangle}{\partial \theta} = \int_{\Omega} \left[\frac{\partial P_{\omega}(t)}{\partial \theta} f_{\omega}(t) + P_{\omega}(t) \frac{\partial f_{\omega}(t)}{\partial \theta} \right] d\Omega \tag{2.2}$$

$$= \int_{\Omega} P_{\omega}(t) \frac{\partial \ln P_{\omega}(t)}{\partial \theta} f_{\omega}(t) d\Omega + \int_{\Omega} P_{\omega}(t) \frac{\partial f_{\omega}(t)}{\partial \theta} d\Omega \tag{2.3}$$

$$= \left\langle \frac{\partial \ln P_{\omega}(t)}{\partial \theta} f_{\omega}(t) \right\rangle + \left\langle \frac{\partial f_{\omega}(t)}{\partial \theta} \right\rangle \tag{2.4}$$

$$= \mathrm{cov} \left(\frac{\partial \ln P_{\omega}(t)}{\partial \theta} f_{\omega}(t) \right) + \left\langle \frac{\partial f_{\omega}(t)}{\partial \theta} \right\rangle \tag{2.5}$$

The last line takes advantage of the fact that $\left\langle \frac{\partial \ln P_{\omega}(t)}{\partial \theta} \right\rangle = 0$ as a consequence of $\int_{\Omega} P_{\omega}(t) = 1$. It is numerically convenient to use covariance rather than expectation, as there is some cancellation of errors. The second term of Equation 2.5 can be sampled directly, but we assume it to be zero.

We estimate the trajectory probability by defining a statistic we term the trajectory derivative (W_{θ}) to track alongside the simulation

$$W_{\theta}(t) = \frac{\partial \ln P_{\omega}(t)}{\partial \theta} \tag{2.6}$$

and therefore compute

$$\frac{\partial \langle f(t) \rangle}{\partial \theta} = \mathrm{cov}\left(W_\theta(t), f(t)\right). \tag{2.7}$$

The statistic $W(t)$ can be computed from simualtion, allowing the right hand side of Equation 2.7 to be computed as an ensemble average. For each trajectory, we may compute $W(t)$ from the reaction propensities a at each KMC step. This is the only modification needed to the standard KMC algorithm in order to perform LRSA. Let n index the steps in the KMC trajectory. i_n indicates the state after n events have occurred. N events occur before time t. $i(t)$ indicates the state of the system at time t and t_n is the time at which the n^{th} event occurs. $a_{i\to j}$ denotes the propensity of the reaction transitioning from state i to state j. We develop the expression[230]

$$W_\theta(t) = \sum_{n=0}^{N-1} \frac{\partial \ln a_{i_n \to i_{n+1}}}{\partial \theta} - \int_0^t \frac{\partial \sum_{j \neq i(t')} a_{i(t')\to j}}{\partial \theta} dt' \tag{2.8}$$

$$= \sum_{i=0}^{N-1} \left[\frac{\partial \ln a_{i_n \to i_{n+1}}}{\partial \theta} - \frac{\partial \sum_{j \neq i_n} a_{i_n \to j}}{\partial \theta}(t_{n+1} - t_n) \right] - \frac{\partial \sum_{j \neq i(t)} a_{i(t)\to j}}{\partial \theta}(t - t_N) \tag{2.9}$$

2.3.1 Equilibrium likelihood ratio

A special case of LRSA method occurs in systems that have a thermodynamic equilibrium. For example, this idea applies to adsorption and desorption of a single species described by an equilibrium isotherm. This approach to computing derivatives has been used widely throughout the Monte Carlo literature, as no kinetic information is needed[146]. In this case, the steady-state probability distribution is described by the Boltzmann distribution and the trajectory derivative $W_\theta(t)$ is not needed. If E is a states dimensionless energy and Q is the partition function, then

$$Q = \sum_i e^{-E_i} \tag{2.10}$$

$$P_i = \frac{e^{-E_i}}{Q} \tag{2.11}$$

Therefore, Equation 2.7 is replaced by

$$\frac{\partial f}{\partial \theta} = \text{cov}\left(-\frac{\partial E}{\partial \theta}, f\right). \tag{2.12}$$

Time dependencies have been dropped under the assumption of steady-state, i.e. it reaches a stationary distribution. We can evaluate the energies during simulation as follows. Take the first state's energy to be zero. Then, each reaction event takes transitions from state i to state j. We start with the equilibrium assumption

$$P_i a_{i\to j} = P_j a_{j\to i} \tag{2.13}$$

$$\frac{P_i}{P_j} = \frac{a_{j\to i}}{a_{i\to j}} \tag{2.14}$$

$$\frac{e^{-E_i}}{e^{-E_j}} = \frac{a_{j\to i}}{a_{i\to j}} \tag{2.15}$$

$$\Delta E_{i\to j} = E_j - E_i = \ln\left(\frac{a_{j\to i}}{a_{i\to j}}\right) \tag{2.16}$$

The propensities are computed during simulation and are used to compute the energies. The advantage of this formaulation is that Metropolis simulation rather than KMC can be used to simulate the equilibrated microscale, because dynamics are not needed and $\Delta E_{i\to j}$ is known.

2.4 Singularly Perturbed KMC Simulations

Singular perturbation is an approximate method for handling large separations of time scales and is widely applied in ordinary differential equations (ODEs)[48]. A system can be reduced at short time scales to the fast eigenmodes by slaving the fast dynamics to the slow ones. The system of slow modes can then be solved at long time scales by taking large timesteps on the macroscale without sacrificing the fast dynamics. Because KMC is governed by a system of linear ordinary differential equations, singular perturbation can be applied there as well.

2.4.1 Theory

The singular perturbation methodology can be applied to the chemical master equation and extended to stochastic systems[205]. Let $s(t)$ be a vector defining the

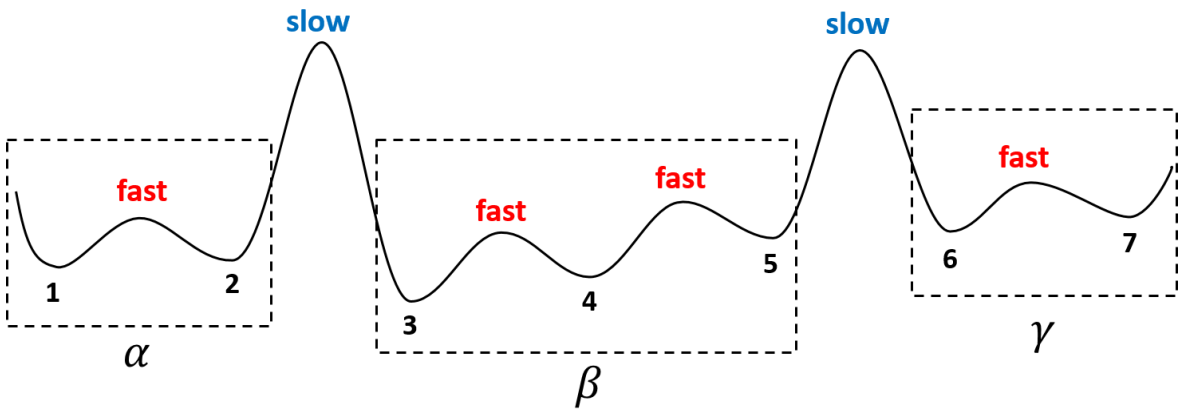


Figure 2.1: Partitioning of the state space. Groups of states which can be linked by fast reactions (low barriers) belong to the same fast class, denoted with Greek letters. Slow reactions (high barriers) separate the fast classes.

state of the system at time t. In well-mixed systems, s lists species populations, whereas in spatial systems, s lists lattice site occupancies. The state space S is the set of all possible states the system can visit. By labeling each reaction as either fast or slow, based on its frequency (for automation of computational singular perturbation method to large networks, see Samant and Vlachos[188]) the state space S is partitioned into disjoint subspaces as follows. States i and j are defined to be members of the same fast class if and only if there exists a sequence of fast reactions which transitions between them (see Figure 2.1). The state space is partitioned as S as $S = \cup_{\alpha\in\Omega} S_\alpha$, where Ω is an indexing set and $S_\alpha \cap S_\beta = \emptyset$ for $\alpha \neq \beta$. We use the notation $a_{i\to j}$ to denote the propensity of the reaction which transitions the system from i to j. The chemical master equation[205, 65] is written as

$$\frac{dp_i(t)}{dt} = \sum_{j\in S} [a_{j\to i}p_j(t) - a_{i\to j}p_i(t)] \tag{2.17}$$

and governs the time evolution of a stochastic system. $p_i(t)$ is the probability that the system is in the state i at time t. $a_{i\to j}$ is the propensity of the reaction transitioning from state i to state j. If $i \in S_\alpha$, then state partitioning[205] allows for the

decomposition of Equation 2.17 into

$$\frac{dp_s(t)}{dt} = \frac{1}{\epsilon}\sum_{j\in S_\alpha}\left[a_{j\to i}p_j(t) - a_{i\to j}p_i(t)\right] + \sum_{j\notin S_\alpha}\left[a_{j\to i}p_j(t) - a_{i\to j}p_i(t)\right]. \tag{2.18}$$

The variable $\epsilon << 1$ captures the time scale separation. At any given time, a fast reaction is much more likely to fire than a slow reaction. It is not always apparent from the propensities where the time scale separation occurs. Samant and Vlachos discuss how to partition the system into the fast and slow modes based on network analysis[187]. The fast time scale can be accessed by taking $\epsilon \to 0$ in Equation 2.18, thus eliminating the time dependence on the left hand side and leaving

$$0 = \sum_{j\in\alpha}\left[a_{j\to i}p_j(t) - a_{i\to j}p_i(t)\right]. \tag{2.19}$$

The solution to Equation 2.19 gives the time-invariant probability distribution function (PDF) for the states within S_α. In simulation, an estimate is obtained by firing only fast reactions, which we will refer to as microscopic integration. We also define

$$p_\alpha(t) = \sum_{i\in S_\alpha} p_i(t) \tag{2.20}$$

to be the total probability of the fast class. The value of $p_\alpha(t)$ is time dependent, but the ratio $\frac{p_i}{p_\alpha}$ is time independent.

The slow time scales can be accessed by substituting the solution of Equation 2.19 into Equation 2.18, giving the time evolution along the slow time scale as

$$\frac{dp_\alpha(t)}{dt} = \sum_{\alpha\in\Omega}\left[a_{\beta\to\alpha}p_\beta(t) - a_{\alpha\to\beta}p_\alpha(t)\right] \tag{2.21}$$

$$a_{\alpha\to\beta} = \sum_{i\in\alpha}\sum_{j\in\beta} a_{i\to j}\frac{p_i}{p_\alpha} \tag{2.22}$$

In Equation 2.21, state probabilities $p_i(t)$ and transition propensities $a_{i\to j}$ are replaced with microscopically averaged fast class probabilities $p_\alpha(t)$ (Equation 2.22) and transition propensities $a_{\alpha\to\beta}$. We use the term microscopic averaging to refer to taking averages during the firing of fast reactions only. Macroscopic averaging refers to time-averages over the slow jumps in the system.

2.4.2 Implementation

We present a simplified version of the algorithm used by Samant and Vlachos[187]. Two nested loops are used. The outer loop executes the macroscopic time steps and the inner loop performs the microscopic averaging. At the beginning of each macroscopic time step, the microscopic loop is run until convergence of the fast class PDF is reached. Convergence is typically measured with either a partial equilibrium criterion[187] or t-test criterion[188]. In this work, for simplicity, we use a fixed number of events for microscale averaging. We treat this as an adjustable parameter, which we tune until good statistics are achieved. Although only fast events fire, slow propensities are also tracked. Subsequently, a slow reaction is chosen to fire. The probability of choosing the slow reaction to fire from state i to state j in a different fast class is

$$p(i \to j) = \frac{p_i a_{i\to j}}{\sum_{k\in S_\alpha} p_k \sum_{m\notin S_\alpha} a_{k\to m}}. \tag{2.23}$$

The macroscopic time clock is advanced by an increment chosen from the exponential distribution with expectation

$$E[\Delta t] = \left(\sum_{k\in S_\alpha} p_k \sum_{m\notin S_\alpha} a_{k\to m} \right)^{-1}. \tag{2.24}$$

2.5 LRSA applied to singularly perturbed KMC simulations

In this section, we discuss how the LRSA appraoch is applied to systems in which disparate time scales have been separated with singular perturbation. An analogous formulation for deterministic simulations is outlined in Section A.1.

2.5.1 Theory

Gupta and Khammash[76] have shown that for two time scale Markov chains we have

$$\lim_{\epsilon\to 0} \frac{\partial \langle f \rangle}{\partial \theta} = \frac{\partial}{\partial \theta} \lim_{\epsilon\to 0} \langle f \rangle . \tag{2.25}$$

In Equation 2.25, the left hand side is the derivative in the single time scale stiff system, while the right hand side is the derivative in the two time scale system. The right hand side can be computed much more reliably than the left hand side, as we will show.

In a two time-scale system, we write the macroscopic average of a property f and take the derivative. Hereafter, an overbar denotes microscopic averaging whereas brackets denote macroscopic averaging.

$$\langle f(t) \rangle = \sum_{\alpha \in \Omega} p_\alpha(t) \bar{f}_\alpha(t) \tag{2.26}$$

$$\frac{\partial \langle f(t) \rangle}{\partial \theta} = \sum_{\alpha \in \Omega} \left[\frac{\partial p_\alpha(t)}{\partial \theta} \bar{f}_\alpha(t) + p_\alpha(t) \frac{\partial \bar{f}_\alpha(t)}{\partial \theta} \right] \tag{2.27}$$

$$= \sum_{\alpha \in \Omega} p_\alpha(t) \frac{\partial \ln p_\alpha(t)}{\partial \theta} \bar{f}_\alpha(t) + \sum_{\alpha \in \Omega} p_\alpha(t) \frac{\partial \bar{f}_\alpha(t)}{\partial \theta} \tag{2.28}$$

$$= \mathrm{cov}\left(\bar{f}(t), W_\theta^M(t) \right) + \left\langle \frac{\partial \bar{f}_\alpha(t)}{\partial \theta} \right\rangle \tag{2.29}$$

The first term on the right hand side of Equation 2.28 involves a derivative of the macroscale trajectory probability and can be written analogously to Equation 2.7. The second term is a simple macroscale average.

Equation 2.29 effectively separates the derivative into contributions from the two time scales. In the first term, the parameter θ enters only in how it affects the macroscale trajectory. In the second, it enters only in how it affects the microscale averages. Thus, the first and second terms capture the effect of θ on the macroscopic and microscopic averaging, respectively. Each one represents a derivative at a specific time scale where stiffness is not present, so a standard method for computing each derivative can be used. Specifically, Equation 2.7 is used for the slow scale contribution and Equation 2.12 is used for the fast scale contribution.

We consider how parameters affecting the slow and fast reactions behave on different time scales. For slow reactions, the computation of the first term does not encounter any problems as far as the time window, over which averages are taken, is appropriate. For fast reactions, the application of Equation 2.7 for computing the first term must be changed. We must measure how fast rate constants affect the macroscopic

trajectory. Although fast rate constants do not affect the instantaneous propensities of slow reactions, they affect the microscopic PDF over which the slow propensities are averaged. Therefore, we define a new statistic we term the multiscale trajectory derivative, analogous to Equation 2.9. n indexes slow events only.

$$W_\theta^M(t) = \frac{\partial \ln p_{\omega,M}(t)}{\partial \theta} \tag{2.30}$$

$$W_\theta^M(t) = \sum_{n=0}^{N-1} \left[\frac{\partial a_{\alpha_n \to \alpha_{n+1}}}{\partial \theta} - \frac{\partial \sum_{\beta \neq \alpha_n} a_{\alpha_n \to \beta}}{\partial \theta} (t_{n+1} - t_n) \right] + \frac{\partial \sum_{\beta \neq \alpha(t)} a_{\alpha(t) \to \beta}}{\partial \theta} (t - t_N) \tag{2.31}$$

Using this additional macroscopic statistic, we can compute the derivative in the same way as we do for the slow rate constants, but using $W_\theta^M(t)$ in place of $W_\theta(t)$. In computing the second term of Equation 2.29, slow reactions have no influence over microscopic averaging, and therefore, $\frac{\partial \bar{f}}{\partial \theta_{slow}} = 0$. We term the above method as multiscale likelihood ratio sensitivity analysis (MLRSA).

2.5.2 Pseudocode

We summarize the proposed method as follows.

1. Initialize the system at $t = 0$.
2. Label all reactions as either fast or slow.

Microscopic Averaging

3. Carry out Metropolis algorithm allowing only fast reactions to fire. Stop when the state PDF, p_α, is converged.
4. Compute microscopic property averages $\bar{f}$ as well as slow propensity averages $a_{\alpha \to \beta}$.

Macroscopic Advancement

5. Select a slow reaction to fire according to Equation 2.23. Update species populations.
6. Update Δt by selecting a time step from exponential distribution according to Equation 2.24.
7. Update $W_\theta^M(t)$ according to Equation 2.31

Table 2.1: Reaction network for the A↔B→C model. Rate constants (r.c.) are in units of inverse seconds. An initial population of $* = 100$ is used.

Reaction	r.c. name	r.c. value	Propensity	Classification
A→B	k_1	$2\epsilon^{-1}$	$k_1 N_A$	Fast
B→A	k_2	$3\epsilon^{-1}$	$k_2 N_B$	Fast
B→C	k_3	2.0	$k_3 N_B$	Slow

8. Terminate the simulation if maximum simulation time has been reached. Otherwise, return to Step 3.

Post-processing

9. Compute simulation property averages using Equation 2.26.

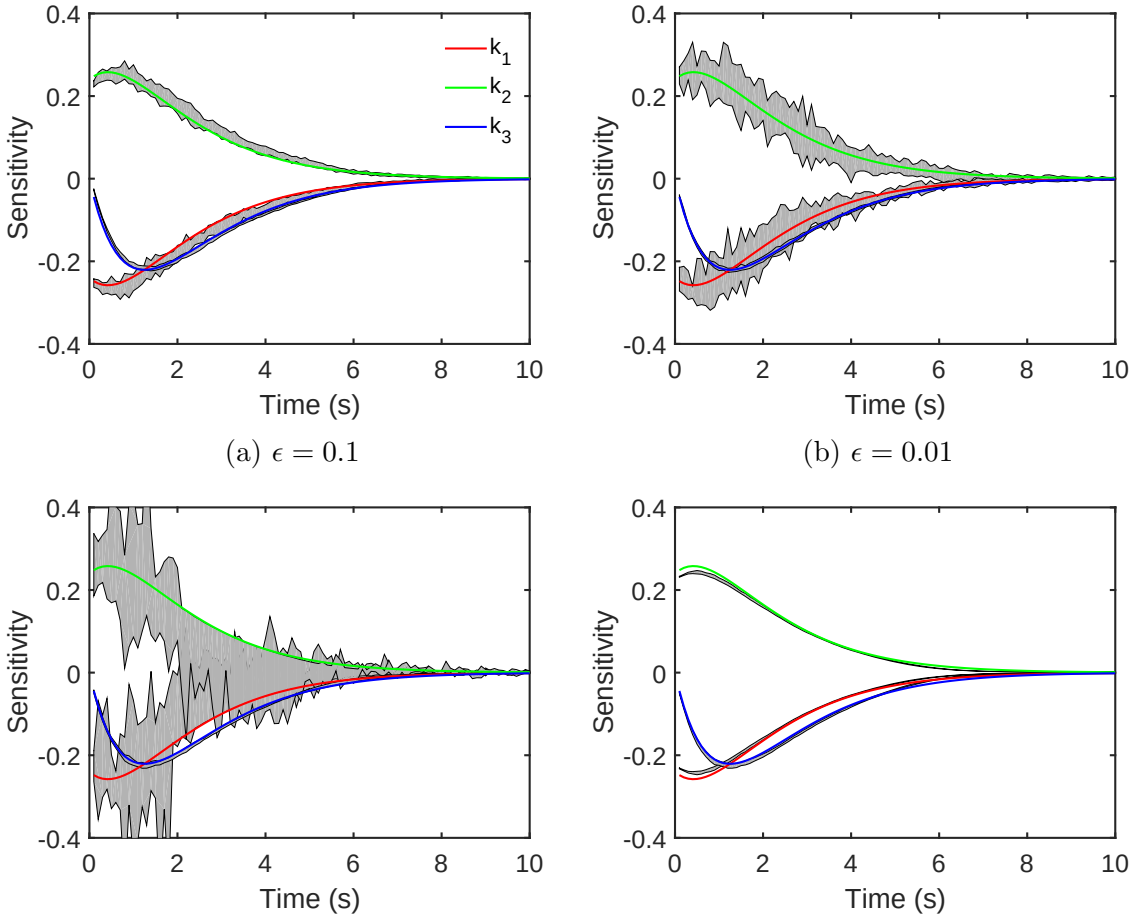
10. Compute sensitivity coefficients using Equation 2.29.

2.5.3 Numerical example

We test our method on a linear system, for which sensitivity coefficients can be computed analytically using formulae from Section A.2. We simulate the system using an open source C++ software we have developed, which is available at `https://github.com/VlachosGroup/Multiscale-KMC`. Table 2.1 shows the details of the reaction network. Reactant species A rapidly interconverts with the intermediate species B, which then forms the product species C. The parameter ϵ is varied to set the stiffness of the system. Notice that the equilibrium constant $K = \frac{k_1}{k_2}$ for the A↔B reaction is constant. Table 2.2 shows the time scales for each level of stiffness as computed through eigenvalue analysis of the chemical master equation (see Section B.1.1). As ϵ decreases, the time scale for the fast reactions also decreases and many more fast reactions fire in between slow reactions. The time scale for the slow reactions is mostly unaffected, so all simulations are run to a termination time of $t_f = 10$s. In this specific example, we set the interconversion of species A and B to be in partial equilibrium as ϵ decreases.

We define our sensitivities as

$$\text{Sensitivity} = \frac{1}{A(0)} \frac{\partial \langle f \rangle}{\partial \ln k}. \tag{2.32}$$



(a) $\epsilon = 0.1$

(b) $\epsilon = 0.01$

(c) $\epsilon = 0.001$

(d) Two time scale.

Figure 2.2: Sensitivity analysis results for different levels of time scale separation.

The property of interest f is taken to be the species populations of A, B, or C. Figure 2.2 shows the results of sensitivity analysis on the model for each level of stiffness. In all cases, 10,000 trajectories are used for averaging. 10 repeated calculations are used to draw error regions of 90% confidence. Using the single time scale LRSA method, the derivative estimates with respect to the rate constants of the slow reactions are accurate for all values of stiffness because slow reactions do not fire often, and thus, they do not accumulate significant noise in the trajectory derivative ($W_k(t)$). However, the derivative estimates with respect to the rate constants for the fast reactions encounter noise larger than the range of physical values as stiffness increases. For very stiff systems ($\epsilon = 0.001$), the sensitivity of species A with respect to k_2 is negative in some of our calculations, whereas it should be positive (Figure 2.2c).This example indicates that the noise in stiff systems may lead to physically wrong conclusions.

This issue is alleviated through the use of the MLRSA approach. Figure 2.2d shows that these estimates are much less noisy than that of the single time scale LRSA approach. Microscopic averaging was performed using 1000 events of equilibration, which was manually adjusted until suitable derivative estimates were achieved. Figure 2.3 shows a parity plot for the calculation of $\frac{\partial \bar{a}_3}{\partial \ln k_1}$ during microscale equilbiration during a single trajectory. The exercise is repeated using 3000 and 10000 microscale events for equilibration, showing that the estimates are mostly converged. Propagation of the error in the microscale sensitivity estimates through Equations 2.31 and 2.29 to estimate the sensitivity is not well understood, but is a possible avenue for future work. For the most part, the noise averages out on the macroscale and is able to provide accurate sensitivity coefficients.

2.6 Conclusions

In this chapter, the likelihood ratio method for efficiently performing sensitivity analysis in stochastic simulations was extended to singularly perturbed systems. Parameter sensitivities were computed separately on the microscopic and macroscopic scales and then summed. The key novelty to this approach is that it accounts for how

Table 2.2: Time scales for the A↔B→C model. Time scales are computed using eigenvalue analysis of the chemical master equation. Total fast events is the average number of fast events that fire in a simulation and is used as a measure of computational cost.

ϵ	τ_{fast}	τ_{slow}	Total fast events
10^0	$1.57 * 10^{-1}$	1.593	401.311
10^{-1}	$1.95 * 10^{-2}$	1.280	3085.18
10^{-2}	$2.00 * 10^{-3}$	1.253	30731.2
10^{-3}	$2.00 * 10^{-4}$	1.250	296008
TTS - 1000	$\to 0$	1.250	10^5
TTS - 3000	$\to 0$	1.250	$3 * 10^5$
TTS - 10000	$\to 0$	1.250	10^6

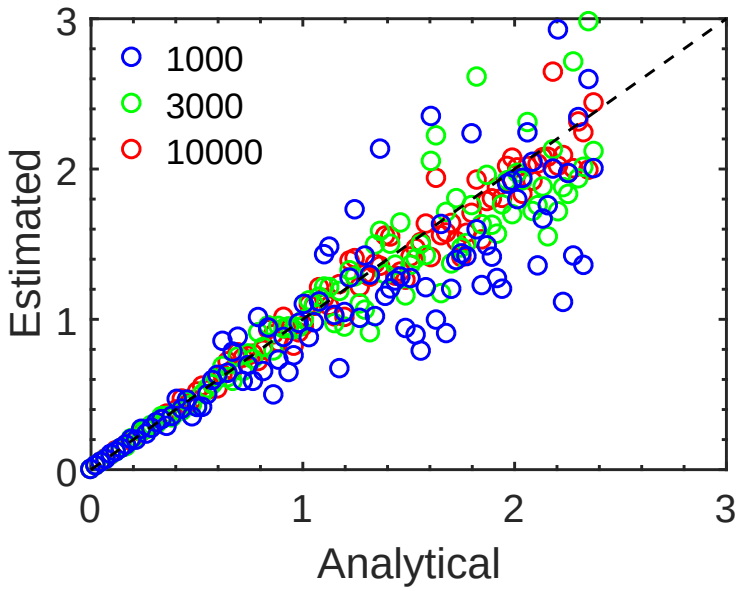


Figure 2.3: Parity plot of microscale derviative estimates computed during a single trajectory. Quantities shown here are $\frac{\partial \bar{a}_3}{\partial \ln(k_1)}$. Colors indicate the number of events used for microscale averaging.

reactions on the fast time scale influence the probability distribution of slow reaction propensities. Parameter influences on fast dynamics can therefore be propagated to the slow time scale. The influence of fast and slow reactions on their respective scales is computed using previously established methods. The proposed technique was applied to an analytical system exhibiting disparate reaction time scales. It was found that the likelihood ratio technique fails at high levels of stiffness due to statistical noise. In contrast, the multiscale likelihood ratio technique computes sensitivities correctly and with little noise. Although in this work, a stochastic simulation algorithm is used for microscopic averaging, the MLRSA technique is generalizable to other multiscale algorithms. For instance, other works have used a deterministic model[81] or the chemical Langevin equation[186] to approximate the microscopic dynamics. As long as the computation of the derivatives in 2.31 can be obtained, sensitivity coefficients can be computed in an analogous way. Thus, MLRSA can be applied to a broad range of problems and approximate methods dealing with stiffness.

Chapter 3

ACCELERATION AND SENSITIVITY ANALYSIS OF LATTICE KINETIC MONTE CARLO SIMULATIONS USING PARALLEL PROCESSING AND RATE CONSTANT RESCALING

3.1 Abstract

Kinetic Monte Carlo (KMC) simulation provides insights into catalytic reactions unobtainable with either experiments or mean field microkinetic models. Sensitivity analysis of KMC models assesses the robustness of the predictions to parametric perturbations and identifies rate-determining steps in a chemical reaction network. Stiffness in the chemical reaction network, a ubiquitous feature, demands lengthy run times for KMC models and renders efficient sensitivity analysis based on the likelihood ratio method unusable. We address the challenge of efficiently conducting KMC simulations and performing accurate sensitivity analysis in systems with unknown time scales by employing two acceleration techniques: rate constant rescaling and parallel processing. We develop statistical criteria that ensure sufficient sampling of non-equilibrium steady state conditions. Our approach provides the twofold benefit of accelerating the simulation itself and enabling likelihood ratio sensitivity analysis, which provides further speedup relative to finite difference sensitivity analysis. As a result, the likelihood ratio method can be applied to real chemistry. We apply our methodology to the water-gas shift reaction on Pt(111).

3.2 Introduction

Operando spatiotemporal understanding at the nanoscale becomes increasingly important but is currently difficult to acquire experimentally. Kinetic Monte Carlo (KMC) is ideal for filling this gap in diverse areas including catalysis, materials growth,

and biological processes. In catalysis, KMC simulations has been applied to many chemistries, such as CO oxidation[164, 202], NO oxidation[234], water gas-shift[201, 237], NH_3 decomposition[73, 75], and methanol decomposition[123] on nanoclusters and heterogeneous surfaces and is suitable for structure sensitive reactions, something for which mean-field microkinetic models are inadequate[205, 206, 213]. Insights gained from KMC models, such as the rate determining step (RDS), dominant reaction pathways, surface coverages, and active sites, can eventually aid catalyst design[183]. Despite the importance of KMC simulation, its widespread application to real chemistries is limited due to a number of challenges discussed next.

Identification of the RDS is particularly important, and requires parametric sensitivity analysis on the rate constants of elementary steps. While sensitivity analysis is routine for mean-field systems[207, 210], the computational burden of KMC simulation prevents sensitivity analysis from becoming common practice in stochastic systems, except for very small networks with a few species and reactions. In larger systems, such as those encountered in heterogeneous catalysis, sensitivity analysis is either performed with an expensive finite difference scheme[75] or not at all. Finite difference techniques require the user to choose an appropriate perturbation size to manage the variance and the bias[133]. The variance can be reduced by using coupled finite difference (CFD)[6, 199], but this approach requires changes to the simulation algorithm, especially for lattice simulations. Furthermore, the number of simulations required still scales linearly with the number of reactions. Likelihood ratio sensitivity analysis (LRSA)[66, 67, 147, 167] is a promising technique for minimizing the computational burden of sensitivity analysis of stochastic systems. Unlike finite difference techniques, it requires only a single (the nominal) run and does not require changes to the underlying simulation algorithm. Due to the challenge of obtaining quality statistical sampling in systems with multiple reaction time scales[155], LRSA has seen only limited application in realistic catalytic systems[87].

Disparate reaction time scales (known also as stiffness) are ubiquitous in real chemistries, plague KMC simulation[41, 206] and introduce enormous noise into LRSA

estimates[155, 230]. Significant work has addressed the issue of stiffness in stochastic systems. One class of multiscale techniques explicitly separates the time scales by treating the fast scales separately and evolves the slow scales on the stochastic manifold of the fast ones[36, 35, 52, 62, 68, 81, 170, 185, 186]. Subsequent work has demonstrated how to apply the LRSA to such systems[76, 82, 155]. Separating a system into fast and slow modes is difficult to do a priori, and the implementation can be burdensome, especially for spatially distributed systems. Single time-scale acceleration techniques include tau leaping[34, 40, 171], net-event algorithms[195, 198], and rescaling the reaction rate constants[42, 47, 51, 57]. Although theoretical advances in this area have been significant, there are few easily usable algorithms for practical systems. Recent work by Hoffman et al. uses a combination of Fisher information matrix, linear response, and coupled finite difference techniques to obtain accurate sensitivity estimates for a CO oxidation model[87]. However, their algorithm requires simulating a lengthy trajectory on a single processor, which would result in significant computational time for stiff systems.

In this chapter, we address the challenges of conducting KMC simulation of real chemistries with multiple time scales and performing accurate sensitivity analysis in systems with unknown time scales by employing two acceleration techniques: rate constant rescaling and parallel processing. Rescaling adjusts the rate constants of fast reactions to considerably reduce the CPU requirement of a KMC trajectory. Parallel processing simulates multiple trajectories simultaneously, so that time and trajectory averaging are combined to enable maximal sampling for a given simulation time. An automated criterion for detecting steady state ensures that sufficient data is sampled to obtain accurate, unsupervised sensitivity estimates using LRSA. The rescaling and steady state detection procedures inform the sensitivity analysis. Rescaled reactions are known to be insensitive, so they are excluded from the LRSA, as their sensitivity estimates would have large variances due to frequent reaction events[230]. Steady state detection identifies the relaxation time scale of the system that is needed for LRSA. The LRSA estimates identify the slow reaction(s) that is(are) RDS. Our analysis

complements existing literature, and is demonstrated for the water-gas shift chemistry. Our use of the LRSA method in real chemistry bridges the gap between mathematical theory and application. Additionally, we provide an open source Python library (available at https://github.com/VlachosGroup/Zacros-Wrapper) which interacts with the Zacros[148, 204] commercial KMC package, and executes the statistical algorithms discussed herein.

3.3 Methods

Our methodology consists of three tasks. First, a robust criterion for detecting steady state is used to ensure that the slow time scale of the system is adequately sampled. Second, the rate constants of equilibrated (fast) reactions are decreased to reduce the CPU without affecting the slow dynamics of the system. Third, the information of Tasks 1 and 2 is used to perform LRSA and identify the RDS.

3.3.1 Steady state identification using combined time and multi-trajectory sampling

The dynamics of a KMC simulation is governed by a continuous-time Markov chain (CTMC) among numerous possible states. At each point in time, there is a finite probability that the system is in any given state. Solving the probability distribution function (PDF) analytically requires enumerating the state space and computing matrix exponentials[93] (see Section B.1), which is infeasible for most practical systems. In practice, the PDF is estimated through statistical sampling, by simulating replicate trajectories. The expected value of a property f at a given time is estimated via

$$E\left[f\left(t\right)\right] \approx \left\langle f\left(t\right)\right\rangle_{n_{\text{traj}}} = \frac{1}{n_{\text{traj}}} \sum_{i=1}^{n_{\text{traj}}} f_i\left(t\right) \tag{3.1}$$

where $f_i\left(t\right)$ is the value of f at time t for trajectory i. In this work, f can be, for example, the instantaneous reaction rate or coverage of a species. Brackets denote an average taken across n_{traj} independent trajectories generated using different random seeds. Under the assumption that a steady state exists, which is the case for lattice

simulations in which the lattice size imposes a finite number of states (see Section B.1), steady state properties are computed by exploiting the ergodic property. That is,

$$f_{ss} = \lim_{t\to\infty} E\left[f\left(t\right)\right] = \lim_{t_{\mathrm{f}}\to\infty} \frac{1}{t_{\mathrm{f}} - t_{\mathrm{cut}}} \int_{t_{\mathrm{cut}}}^{t_{\mathrm{f}}} f\left(t\right) dt \tag{3.2}$$

where $t_{\mathrm{cut}} \geq 0$ is the cutoff time before steady state is reached (for which data is omitted) and t_{f} is the termination time. For ease of notation, we denote time averages as

$$\left[f\right]_{t_1,t_2} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f\left(t\right) dt. \tag{3.3}$$

Under steady state conditions, the time average in Equation 3.3 functionally depends only on $t_2 - t_1$, so we equivalently write $\left[f\right]_{t_1,t_2} = \left[f\right]_{t_2-t_1}$.

We obtain efficient sampling of f_{ss} by using not only the ergodic property of KMC, but also replicate trajectories, so that f_{ss} is estimated as

$$f_{ss} \approx \left\langle \left[f\right]_{t_{\mathrm{cut}},t_{\mathrm{f}}} \right\rangle_{n_{\mathrm{traj}}}. \tag{3.4}$$

All trajectories are simulated simultaneously on n_{traj} processors to reduce real execution time. When $n_{\mathrm{traj}} = 1$, averaging is purely based on time averaging on a single processor, as is common practice for steady state simulations. When n_{traj} is large, each trajectory needs only to exceed the relaxation time to reach steady state, and statistics is computed by averaging across trajectories.

Convergence to steady state and adequate sampling need to be evaluated. Toward this goal, first the autocorrelation function is approximated as an exponential decay

$$\hat{c}\left(\Delta t\right) = \frac{\mathrm{cov}\left(f\left(t\right), f\left(t + \Delta t\right)\right)}{\mathrm{var}\left(f\right)} \approx \exp\left(-\frac{\Delta t}{\tau_{\mathrm{exp}}}\right) \tag{3.5}$$

where τ_{exp} is the exponential relaxation time obtained from eigenvalue analysis of the generator of the chemical master equation (see Section B.1.1). The variance of the estimator in Equation 3.4 under steady state is (see Section B.1.2)

$$\frac{\mathrm{var}\left(\left\langle \left[f\right]_{t_{\mathrm{cut}},t_{\mathrm{f}}} \right\rangle_{n_{\mathrm{traj}}}\right)}{\mathrm{var}\left(f\right)} \approx \frac{1}{n_{\mathrm{traj}}} * \frac{2\left(\tilde{t} + e^{-\tilde{t}} - 1\right)}{\tilde{t}^2}. \tag{3.6}$$

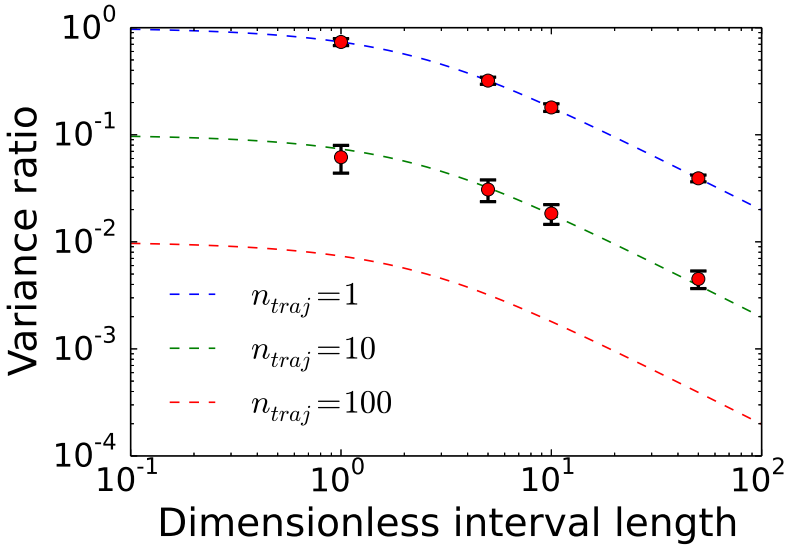


Figure 3.1: Ratio of the variance of the estimate of the mean of a property and the variance of the property itself (i.e., the left hand side of Equation 3.6) as a function of the length of the interval used for averaging ($\tilde{t}$) and the number of replicate trajectories (n_{traj}). Data points show the variances of reaction rates. Simulation details are in Section B.2.1

Time is normalized as $\tilde{t} = \frac{t_{\text{f}} - t_{\text{cut}}}{\tau_{\text{exp}}}$. The significance of Equation 3.6 is that sampling is improved by either increasing the length of each trajectory (i.e., $\tilde{t}$) or increasing the number of replicate trajectories (i.e., n_{traj}) as computational resources allow. Figure 3.1 visualizes Equation 3.6 for several values of n_{traj}. We numerically validate Equation 3.6 by computing average reaction rates for a simple model (see Section 3.4.1) using several values of $\tilde{t}$ and n_{traj}, and then computing the variance of those estimates. Figure 3.1 shows, for example, that the variance of an estimate with 10 trajectories is about the same as that of an estimate using a single trajectory that is 10 times as long. Throughout the rest of this paper, we adopt the strategy of using as many processors as are available to simulate replicate trajectories and gradually increasing the time interval until the variance in Equation 3.6 decays to an acceptable level.

For a simulation, the values of t_{cut} and t_{f} must be chosen. The value of t_{cut} should exceed the decorrelation time of the system. In this work, suitable values of t_{cut} and t_{f} are chosen by gradually increasing the trajectory length (t_{f}) and determining whether

steady state has been achieved using a batch means test[82, 129]. t_f is increased using an iterative approach wherein sets of KMC trajectories are run and data is appended between iterations. Data analysis for an iteration uses the cumulative data from all previous iterations. The final states of the trajectories from the previous iteration are used as the initial states of the trajectories in the next iteration. The final time (t_f) of the cumulative trajectories is increased as

$$t_{\mathrm{f}} = \sum_{i=1}^{n_{\mathrm{iter}}} t_{\mathrm{f}}^{i} \tag{3.7}$$

$$t_{\mathrm{f}}^{i} = s * t_{\mathrm{f}}^{i-1}. \tag{3.8}$$

where t_{f}^{i} is the length of the trajectories in iteration i.

In each successive iteration, the trajectory length is scaled by a factor s; here we use a value of $s = 3$. A fixed number of time points are recorded for each iteration on evenly spaced time intervals, as the time resolution ultimately desired is unknown and intermediate analysis must be performed. As a result, the data storage requirement for the iterative procedure scales with the number of iterations. Smaller values of s require more iterations and therefore record more data on disk. Large values of s risk overshooting suitable values of t_f and simulating longer than necessary. Therefore, s is an adjustable numerical parameter that gives a tradeoff between memory and CPU requirements.

An appropriate value of t_f^1 cannot be determined because the time scales of the system are not known a priori. Therefore, a 1000 event cutoff is used for each trajectory in the first iteration only. The event cutoff must be large enough to estimate the fast time scale, as determined by the time horizon reached within 1000 events, and determine whether rescaling reaction rates in necessary using the event counting method outlined in Section B.2.2. The performance of the algorithm is otherwise insensitive to the event cutoff for the first iteration.

The batch means stopping criterion determines the value of n_{iters}, and therefore t_{f}, at which the iterative process terminates. Following each iteration, each trajectory is partitioned into a fixed number (n_{bpt}) of batches each of length $\Delta t_{\mathrm{batch}} = \frac{t_{\mathrm{f}}}{n_{\mathrm{bpt}}}$,

where batch j spans from t_{j-1} to t_j, and $t_j = j * \Delta t_{\text{batch}}$. The batch length increases proportionally with the trajectory length. When the batch length is sufficiently large, the averages of data in adjacent batches are uncorrelated, that is

$$\text{cov}\left([f]_{t_{j-1},t_j}, [f]_{t_j,t_{j+1}}\right) \ll \text{var}\left([f]_{\Delta t_{\text{batch}}}\right). \tag{3.9}$$

Section B.1.3 derives an analytical expression of how the left hand side of Equation 3.9 should decay. When the batch length is long enough so that Equation 3.9 is satisfied, the average of the second batch, $[f]_{t_1,t_2}$, does not depend on the initial state of the trajectory. Therefore, for the evaluation of Equation 3.2, we use $t_{\text{cut}} = \Delta t_{\text{batch}} = t_1$, so that data in the first batch contains the transient regime that is excluded from statistics.

The batch means approach is modified to account for multiple trajectories. The number of batches per trajectory (n_{bpt}) is chosen so that $n_{\text{bpt}} \geq 3$ and

$$n_{\text{batches}} = n_{\text{traj}} * (n_{\text{bpt}} - 1) \tag{3.10}$$

exceeds a predetermined number. n_{batches} estimates the number of independent and identically distributed data points obtained at steady state. The batch mean for the jth batch of the ith trajectory ($f_{i,j}$) is defined by $f_{i,j} = [f_i]_{t_{j-1},t_j}$. The resulting batch means are tabulated in a $n_{\text{traj}} \times n_{\text{bpt}}$ matrix. Steady state is tested by computing the autocorrelation function (ACF), which is small when the batch length Δt_{batch} exceeds the decorrelation time. The autocorrelation between adjacent batch means is computed by generating two overlapping subsets of the batch means (B_1 and B_2), each offset by the length of a single batch. The autocorrelation function is computed as follows, with Equation 3.13 being the decorrelation criterion.

$$B_1 = \bigcup_{i=1}^{n_{\text{traj}}} \bigcup_{j=2}^{n_{\text{bpt}}-1} f_{i,j} \tag{3.11}$$

$$B_2 = \bigcup_{i=1}^{n_{\text{traj}}} \bigcup_{j=3}^{n_{\text{bpt}}} f_{i,j} \tag{3.12}$$

$$\left|\text{ACF}\left(\Delta t_{\text{batch}}\right)\right| = \left|\frac{\text{cov}\left(B_1, B_2\right)}{\text{var}\left(B_1 \cup B_2\right)}\right| < 0.05 \tag{3.13}$$

The ACF is normalized by an estimator of the steady state value of var (f) and can have an error as it is estimated in a region that may not be at steady state. Therefore, Equation 3.13 is a necessary but not sufficient determination of steady state.

In order to make the steady state convergence more robust, an additional criterion is used. Specifically, f_{ss} should be known to within 5% certainty. The usual t-test statistic is used to construct a confidence interval around the mean. The criterion is

$$\frac{t\left(0.05, n_{\text{batches}} - 1\right) \sqrt{\text{var}\left(\text{mean}\left(B_1 \cup B_2\right)\right)}}{\text{mean}\left(B_1 \cup B_2\right)} < 0.05. \tag{3.14}$$

Here, $t\left(0.05, n_{\text{batches}} - 1\right)$ is the t-test statistic for a 95% confidence interval with $n_{\text{batches}} - 1$ degrees of freedom. The term var $(\text{mean}\left(B_1 \cup B_2\right))$ accounts for correlations between the batch means as detailed in Section B.1.4. Together, Equations 13 and 14 give the criteria for having achieved steady state.

3.3.2 Rescaling of rate constants

Separation of time scales challenges practical simulation, as most of the CPU time is spent on fast, equilibrated elementary steps, which do not advance the slow dynamics of the system. To accelerate simulation, we apply a rate constant rescaling technique following each iteration of the algorithm (for simplicity in this section, we drop off the index indicating an iteration). Rate constants are adjusted with error control on the slow dynamics, as discussed by Yin and Zhang[238], so that the evolution of the system on large time scales and its steady state properties are accurately estimated.

Each elementary step i has a forward rate constant k_i, a reverse rate constant k_{-i}, and an equilibrium constant $K_i = \frac{k_i}{k_{-i}}$. At each iteration of the algorithm, the rate constant of each fast, equilibrated elementary step is scaled down by a factor $\alpha_i \geq 1$.

$$k_i^{\text{new}} = \frac{k_i^{\text{old}}}{\alpha_i} \tag{3.15}$$

$$k_{-i}^{\text{new}} = \frac{k_{-i}^{\text{old}}}{\alpha_i} \tag{3.16}$$

The forward and reverse rate constants are rescaled proportionally so that the equilibrium constants, and thus the thermodynamic properties, are preserved[138]. In essence, this rescaling implies that transition states rather than thermodynamic states change, and thus, the thermochemistry of the entire network is preserved. α_i varies with iteration and is determined based on the frequency of elementary step-firings, as elaborated below.

We determine whether an elementary step is in quasi-equilibrium using the partial equilibrium (PE) index[128, 140, 187]

$$\mathrm{PE}_i = \frac{F_{i,\mathrm{fwd}}}{F_{i,\mathrm{fwd}} + F_{i,\mathrm{bwd}}}. \tag{3.17}$$

Here $F_{i,\mathrm{fwd}}$ and $F_{i,\mathrm{bwd}}$ are the number of times the forward and backwards reactions are fired, respectively, in an iteration of the algorithm. A value of $\mathrm{PE}_i = 0.5$ indicates an equilibrated elementary step, whereas 0 or 1 indicates an irreversible elementary step. If $F_{i,\mathrm{fwd}} = F_{i,\mathrm{bwd}} = 0$, i.e., the elementary step is considered irreversible. For simulation, an elementary step is classified as partially equilibrated if $|\mathrm{PE_i} - 0.5| < 0.05$ and irreversible otherwise. A tolerance of 0.05 has been found to be adequate to account for imperfect equilibration and stochastic fluctuations in prior work[187].

We define a slow scale frequency, F_{slow}, to be the fastest irreversible step, defined by

$$F_{\mathrm{slow}} = \max_{i=\mathrm{irrev.}} \left(1, F_{i,\mathrm{fwd}}, F_{i,\mathrm{bwd}}\right). \tag{3.18}$$

Then for each equilibrated elementary step, we compute the rescaling factor α_i in each iteration according to

$$N_i = \frac{\frac{1}{2}\left(F_{i,\mathrm{fwd}} + F_{i,\mathrm{bwd}}\right)}{F_{\mathrm{slow}}} \tag{3.19}$$

$$\frac{N_i}{N_{\mathrm{min}}} \geq \alpha_i. \tag{3.20}$$

Here N_i is the average number of firings of elementary step i per slow event. N_{min} places an upper bound on how much a rate constant can be rescaled without affecting the slow dynamics. A value $N_{\mathrm{min}} = 50$ was chosen to balance accuracy and computational cost, as detailed in Section B.2.2.

3.3.3 Likelihood ratio sensitivity analysis (LRSA)

Sensitivity analysis is key for complex systems and is typically carried out using finite difference. For our application, f is the rate of a catalytic reaction, as computed from reaction frequencies and stoichiometry. Other quantities of interest may also be estimated using the same approach. The net gas-phase reaction can be written in stoichiometric form. For example, the water-gas shift reaction is written as

$$\mathrm{CO} + \mathrm{H_2O} \leftrightarrow \mathrm{CO_2} + \mathrm{H_2}. \tag{3.21}$$

To compute the rate, we pick a product species P and take its stoichiometric coefficient ν_P. In the water-gas shift reaction, P is CO_2 and $\nu_P = 1$. The reaction rate $r(t)$ at any given time is the rate of change of the product species population $P(t)$ normalized by its stoichiometric coefficient and the number of lattice sites (n_{sites}).

$$r(t) = \frac{1}{n_{\text{sites}}\nu_P}\frac{dE[P(t)]}{dt} \tag{3.22}$$

In simulation, $\frac{dE[P(t)]}{dt}$ is equal in expectation to the sum of the propensities of all reactions that produce P, so

$$r(t) = \frac{1}{n_{\text{sites}}\nu_P}\sum_i \nu_{i,P} E[a_i(t)] \tag{3.23}$$

where $\upsilon_{i,P}$ is the stoichiometric coefficient of P in elementary step i and $a_i(t)$ is the propensity (reaction rate) of reaction i. The steady state value of the rate, r_{ss}, is computed using Equation 3.4.

The degree of rate control (DRC)[32, 137, 207, 233] for each elementary reaction is defined as

$$\mathrm{DRC}_i = \left(\frac{\partial \ln r_{ss}}{\partial \ln k_i}\right)_{K_i, k_{j\neq i}}. \tag{3.24}$$

Sensitivity analysis is performed as post-processing (see Figure 3.2). For reactions that are partially equilibrated (see Equation 3.17), the overall rate is insensitive to these rate constants and the DRC is set to zero. This simple approach employing partial equilibrium information can save significant computational time when a finite

difference approach is employed. For non-equilibrated reactions, LRSA is used. An essential component to the LRSA is the trajectory derivative (W) for each parameter. Trajectory derivatives are computed through post-processing according to the equation derived in Warren and Allen[230].

$$W_{\ln k_i}(t) = F_i(t) - \int_0^t a_i(t')dt' \tag{3.25}$$

$F_i(t)$ is the number of times the elementary step i has fired up to time t. The integral on the right hand side is the cumulative propensity, which is the expected number of times that reaction i fires between time 0 and time t. An intuitive interpretation of $W_{\ln k_i}(t)$ is that it quantifies the difference between the observed and expected number of firings. $W_{\ln k_i}(t)$ has expectation zero at all times and is only nonzero due to stochastic noise. Sensitivities are computed as the covariance between the trajectory derivative and the rate. A positive covariance indicates that more frequent firings of the elementary step result in a higher observed rate. Sensitivities are estimated by either the centered likelihood ratio (CLR) or the centered ergodic likelihood ratio (CELR) methods, shown respectively in the following two equations.

$$\mathrm{DRC}_i^{\mathrm{CLR}} = \mathrm{cov}\left(W_{\ln k_i}(t) - W_{\ln k_i}(t-\Delta t), \frac{r(t)}{r_{ss}}\right) + \frac{\upsilon_{i,P} a_{i,ss}}{r_{ss}} \tag{3.26}$$

$$\mathrm{DRC}_i^{\mathrm{CELR}} = \mathrm{cov}\left(W_{\ln k_i}(t) - W_{\ln k_i}(t-\Delta t), \frac{[r]_{t-\Delta t,t}}{r_{ss}}\right) + \frac{\upsilon_{i,P} a_{i,ss}}{r_{ss}} \tag{3.27}$$

$a_{i,ss}$ is the steady state value of the propensity of elementary step i, as estimated by Equation 3.4. $\frac{\upsilon_{i,P} a_{i,ss}}{r_{ss}}$ is the fractional contribution of reaction i to the steady state rate in Equation 3.23. Covariance statistics is computed by combining trajectory and time averaging as

$$\mathrm{cov}(x,y) = \left\langle [xy]_{t_1,t_2} \right\rangle - \left\langle [x]_{t_1,t_2} \right\rangle \left\langle [y]_{t_1,t_2} \right\rangle. \tag{3.28}$$

CLR and CELR require a time window Δt. We take advantage of the time scale information gained from the steady state convergence and rate constant rescaling to determine an appropriate value of Δt, the choice of which heavily impacts the

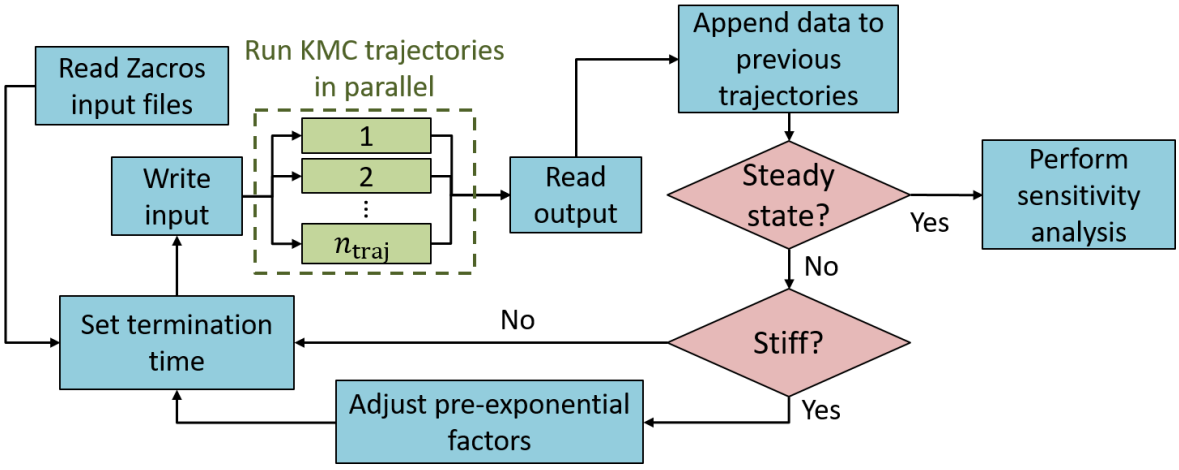


Figure 3.2: Flow chart of the iterative algorithm, detailed in Section 3.3.

performance of these methods. The decorrelation time is estimated as the batch length Δt_{batch} determined in the steady state convergence task. The steady state LRSA technique[230] also requires that Δt exceeds the decorrelation time, so we use $\Delta t = \Delta t_{\text{batch}}$.

The performances of the CLR and CELR methods have been compared in other works[8, 82]. CLR results in noisy estimates (high variance) of the rate. CELR averages the rate over the batch length, resulting in a lower variance but larger bias. A larger minimum batch length must be used to eliminate the bias, but beyond this minimum, the estimate is insensitive to the batch length.

3.3.4 Pseudocode

We show pseudocode for the methods outlined in Section 3.3.1 through 3.3.3.

1. Define the reactions, species, and energetics in Zacros input files. Compute rate constants from ab initio data.
2. Choose n_{bpt} to fit the desired sampling and n_{traj} depending on the number of processors available.
3. Set the maximum number of events equal to 1000.
4. Run n_{traj} trajectories in parallel with different random seeds.
5. Average the event frequencies from all trajectories.

6. Compute α_i for each reaction and rescale rate constants according to Equations 3.15-3.20.
7. Record the average termination time (t_{f}^1) for the trajectories.
8. While Equations 3.13 and 3.14 are not satisfied
    a. Set t_{f}^i using Equation 3.8.
    b. Set up n_{traj} trajectories in parallel. Use the final states of previous iteration's trajectories 1, 2, ..., n_{traj} as the initial states of the new trajectories 1, 2, ..., n_{traj} respectively.
    c. Run the new trajectories for a length of t_{f}^i.
    d. Average the event frequencies from all new trajectories.
    e. Compute α_i for each reaction and rescale rate constants according to Equations 3.15-3.20.
    f. Increment the time indices of the new trajectories by the sum of all previous t_{f}^i. Append the time series data from the new trajectories $1, 2, \ldots, n_{\mathrm{traj}}$ onto the time series data of the cumulative trajectories $1, 2, \ldots, n_{\mathrm{traj}}$ respectively.
9. Perform sensitivity analysis using Equations 3.26 and 3.27.

3.4 Results

The methods developed in the previous section are implemented in a Python interface for the Zacros graph theoretical KMC software[148, 204] and are applied to two examples. First, a prototype model is benchmarked against its analytical solution. Second, the water-gas shift system[201], which is a much more complex system, is simulated, and the results of the sensitivity analysis are physically interpreted.

3.4.1 A prototype A $\rightarrow$ B model

Here we use the simple model, used in previous work[82, 155], of a catalytic reaction with two time scales. The reactant species A adsorbs onto the catalyst surface, isomerizes to species B, and then desorbs to form B in the gas phase. The reactions and rate constants are shown in Table 3.1. That lattice consists of $n_{\mathrm{sites}} = 100$ decoupled sites. Adsorption of A has a significantly faster time scale than the other two reactions. The system is sufficiently simple so that the chemical master equation (CME) can be solved directly, as detailed Section B.2.1.

Table 3.1: A→B model. $P_A = 1$ atm and $P_B = 0$ atm. $K = \frac{k_{\text{fwd}}}{k_{\text{rev}}}$. The surface species are A*, B*, and empty sites (*).

Reaction	Forward rate constant, k_{fwd}	Equilibrium constant K
* + A(g) ↔ A*	$k_1 = 10^5$ atm$^{-1}s^{-1}$	$K_1 = 0.67$ atm^{-1}
A* ↔ B*	$k_2 = 2\ s^{-1}$	$K_2 = 2$
B* ↔ B(g) + *	$k_3 = 0.4\ s^{-1}$	$K_3 = 40$ atm

For numerical simulation, $N_{\text{traj}} = 96$ replicate trajectories are run, each on its own processor. For each trajectory in the first iteration, an empty lattice is used as the initial state. $N_{\text{bpt}} = 12$ batches per trajectory are used for data analysis, so that $n_{\text{batches}} = 1152$ according to Equation 3.10, offering a satisfactory number of statistically independent steady state observations. The convergence of the rate estimate with increasing batch length is shown in Figure 3.3. The iteration number is labeled for each point, demonstrating how the batch length (Δt_{batch}) increases between successive iterations. As the batch length increases, the first batch includes more of the transient period and thus removes more of this bias from the steady state rate estimate. At early times, the lattice is bare, the overall rate is low, and the steady state rate is underestimated. The variance in the rate is small, so the confidence interval on the rate is small for every iteration.

The decay of the autocorrelation between adjacent batches is displayed in Figure 3.4 and compared to an exponential decay function with time scale $\tau_{\text{exp}} = 0.4624\ s$, as computed from the chemical master equation. The autocorrelation between batches is larger than the autocorrelation of instantaneous rates. To account for the difference between batch averages and instantaneous values, we derive an analytical expression for the autocorrelation between batches (see Section B.1.3) shown in the green curve. It more closely matches the simulation data for large batch lengths near steady state. The autocorrelation of the batches converges near zero at iteration 9, so data up to and including iteration 9 is used for sensitivity analysis. The batch length at iteration 9 is $\Delta t_{\text{batch}} = 8.01\ s$, which is 17 times larger than τ_{exp}. Although this batch length greatly exceeds the decorrelation time, which should only be a 3-5 times τ_{exp}, overestimating

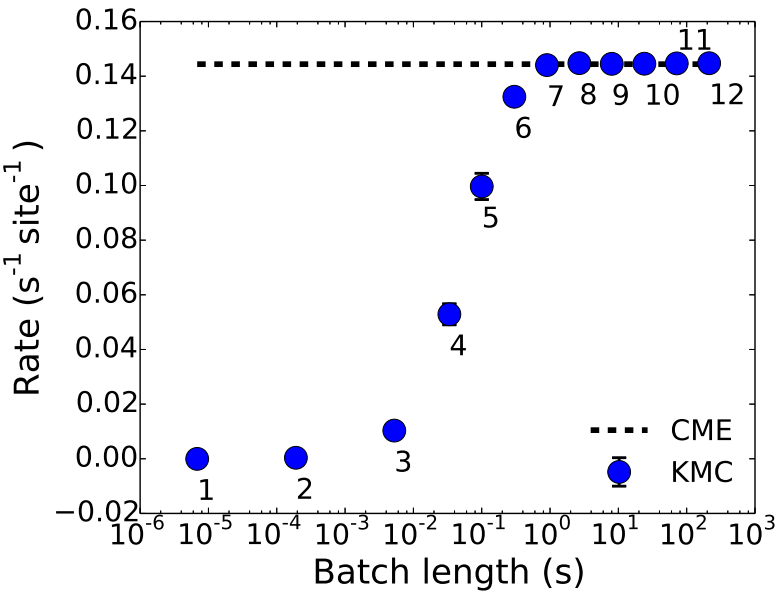


Figure 3.3: Estimates of the mean rate versus batch length in the A→B model. The dashed line is the analytical value obtained from solving the chemical master equation (CME). The points are labeled with the iteration number of the algorithm.

the batch length is advantageous for CELR sensitivity analysis, because a batch length larger than the decorrelation time eliminates bias.

Rate constants are rescaled at each iteration of the algorithm, as shown in Figure 3.5. Adsorption of A is the only reaction that is fast and equilibrated, and therefore rescaled. CELR and CLR sensitivity analyses are performed and the results are shown in Figure 3.6. Both methods predict the sensitivities accurately. The isomerization reaction and desorption of B are kinetically relevant, with the latter being the key rate-determining step. The CLR estimate has high variance due to the batch length being larger than the decorrelation time. In contrast, the CELR estimate is both accurate and of low variance.

3.4.2 Water-gas shift (WGS) reaction

The model for the WGS reaction on Pt(111) developed by Stamatakis et al.[201] has also been evaluated. The WGS reaction network contains 19 elementary reactions and 8 surface species. The Pt(111) lattice consists of 200 top sites, 600 bridge sites, 200

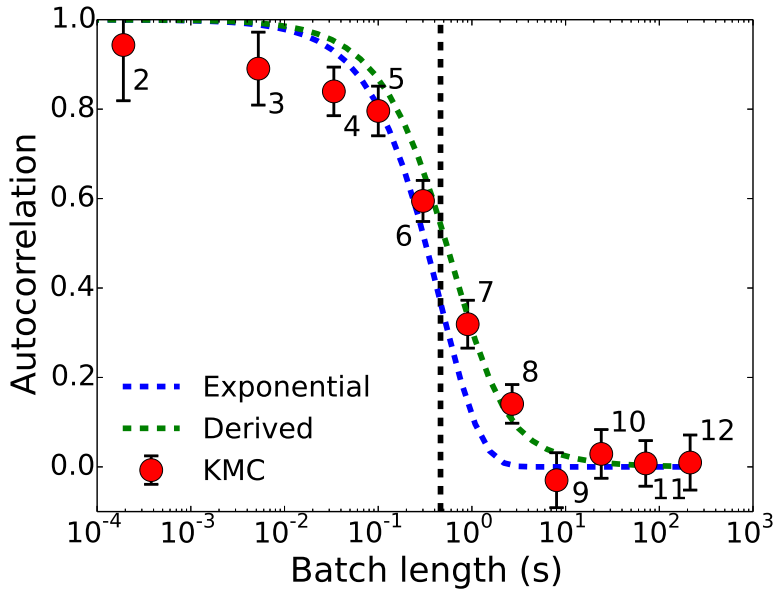


Figure 3.4: Computed values for the autocorrelation as a function of batch length in the A→B model. The blue line is the exponential decay for the characteristic time scale obtained from eigenvalue analysis of the chemical master equation. The vertical black line indicates the exponential autocorrelation time. The green line is a decay function that accounts for batch averaging, and is derived in Section B.1.3

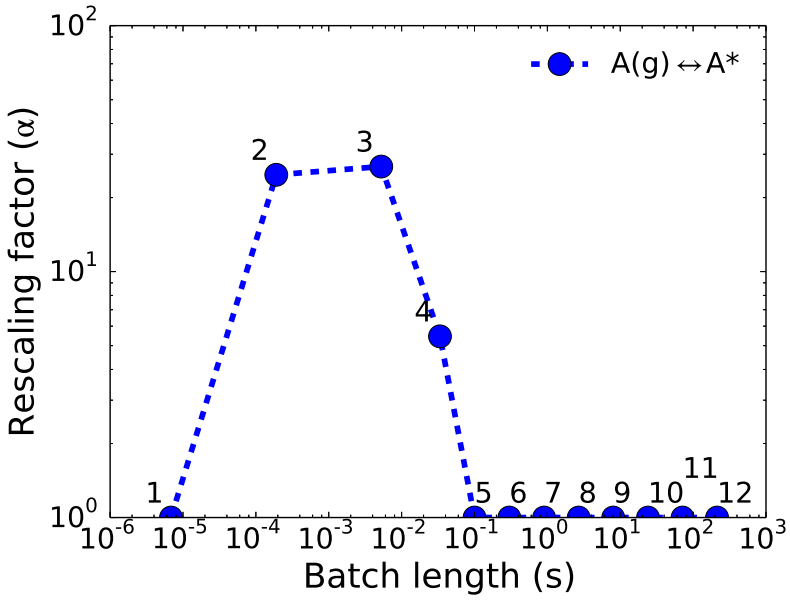


Figure 3.5: Rate-constant rescaling factors used in the A→B model. Adsorption of A is the only fast reaction. Once adsorption of A has been rescaled enough to remove stiffness, no more rescaling is done. The points are labeled with the iteration number of the algorithm.

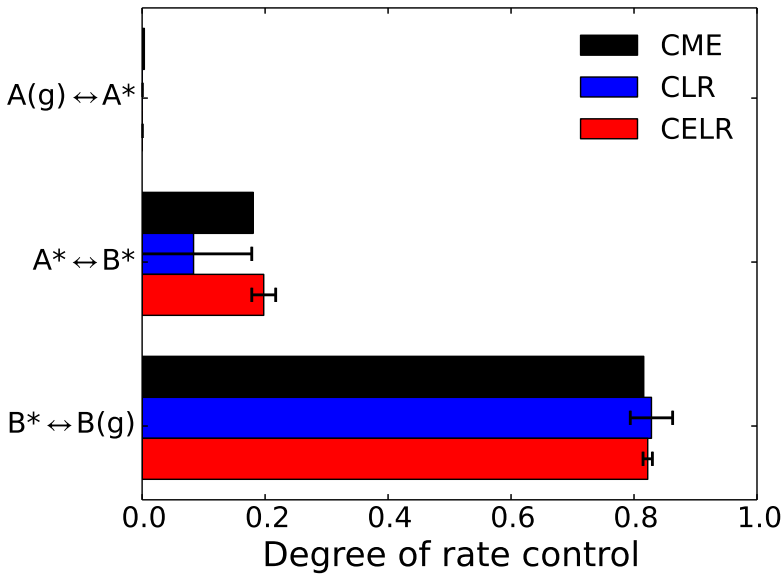


Figure 3.6: Sensitivity analysis for the A→B model. The degree of rate control as a measure of sensitivity is defined in Equation 3.24. Estimates are computed using the centered likelihood ratio (CLR) (blue) and centered ergodic likelihood ratio (CELR) (red) methods. Analytical values derived from the chemical master equation (CME) (black) are used as a benchmark.

fcc hollow sites, and 200 hcp hollow sites. Consistent with convention, we normalize the reaction rate by the number of top sites, so that $n_{\text{sites}} = 200$. We initialize the simulation with the rate constants as evaluated in the original work, which used density functional theory calculations. Values of $N_{traj} = 96$ and $N_{bpt} = 12$ are used for the steady state convergence. The convergence of the rate estimates and autocorrelation are shown in Figure 3.7 and Figure 3.8, respectively. In contrast to the A→B model, the variance of the rate is very large due to short-lived states with high propensities contributing to the rate in Equation 3.23. The error bar on the rate estimate only decays below the allowable threshold at iteration 7. The autocorrelation estimates are at first noisy and decrease after iteration 5, but are low enough to rule out correlations between batches. Therefore, data up to and including iteration 7 (excluding the transient region) is used for the sensitivity analysis. In Stamatakis et al.[201], the authors manually decreased the rate constants of CO and H_2O adsorption by two orders of magnitude. We use the nominal (unadjusted) rate constants and allow the rescaling algorithm to adjust the rate constants in an unsupervised manner. The results of this process are shown in Figure 3.9. The rate constant for H_2O adsorption is decreased by over three orders of magnitude without affecting the slow dynamics. Section B.3 shows how elementary step frequencies are affected by the rescaling.

The sensitivities of the rate to each of the rate constants, as estimated by the finite difference and CELR methods, are shown in Figure 3.10. Only H_2O* dissociation and CO*+OH* disproportionation have significant sensitivities, with H_2O* dissociation being the primary RDS. The variance of the CLR estimates is so large that it yields unphysical values of the DRC and is not shown. Finite difference sensitivity analysis uses a 5% perturbation size and no coupling. The finite difference and CELR methods agree within statistical error, thus validating performance of the CELR. The variance of the CELR estimates, as estimated with statistical bootstrapping, is relatively small for the rate determining steps. The rate determining steps occur infrequently during simulation, thus avoiding high variance. The variance of the finite difference method would be expected to be much lower if coupling were used, but is small enough to show

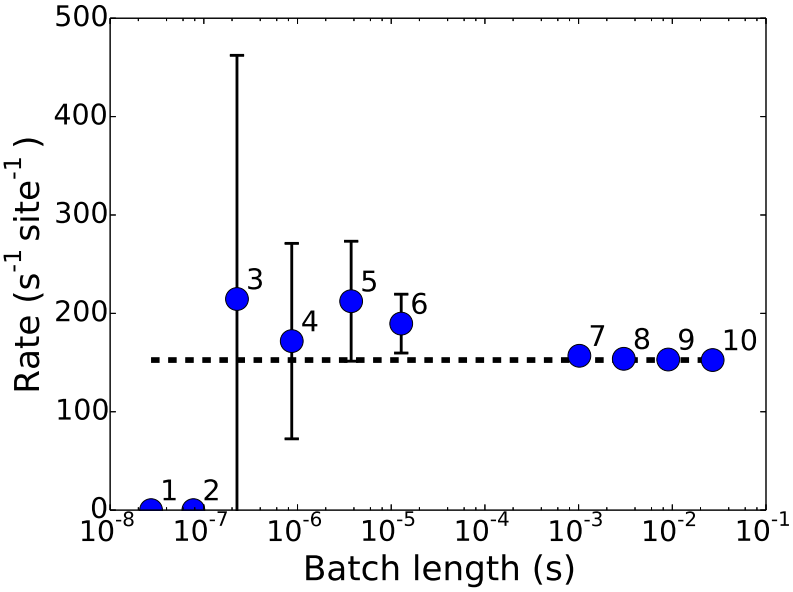


Figure 3.7: Estimates of the mean rate versus batch length in the WGS model. The dashed, horizontal line indicates the final, most accurate estimate. The points are labeled with the iteration number.

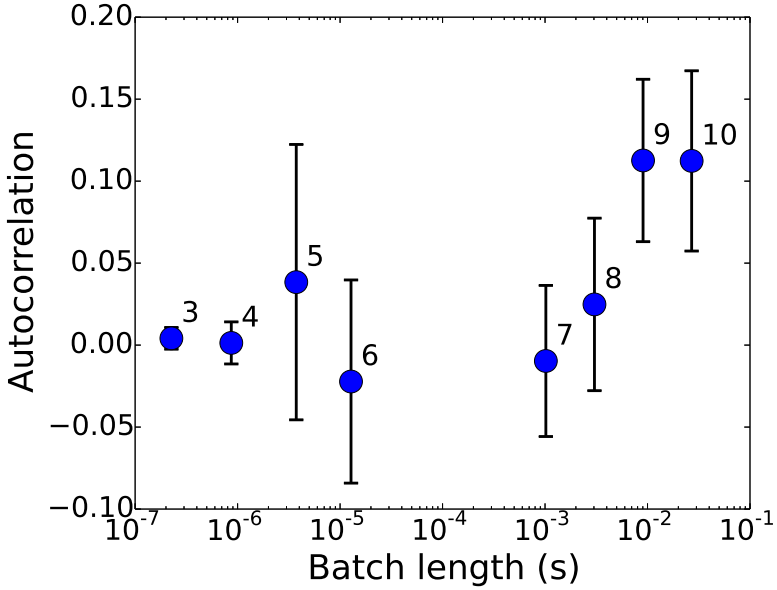


Figure 3.8: Computed values for the autocorrelation as a function of batch length in the WGS model.

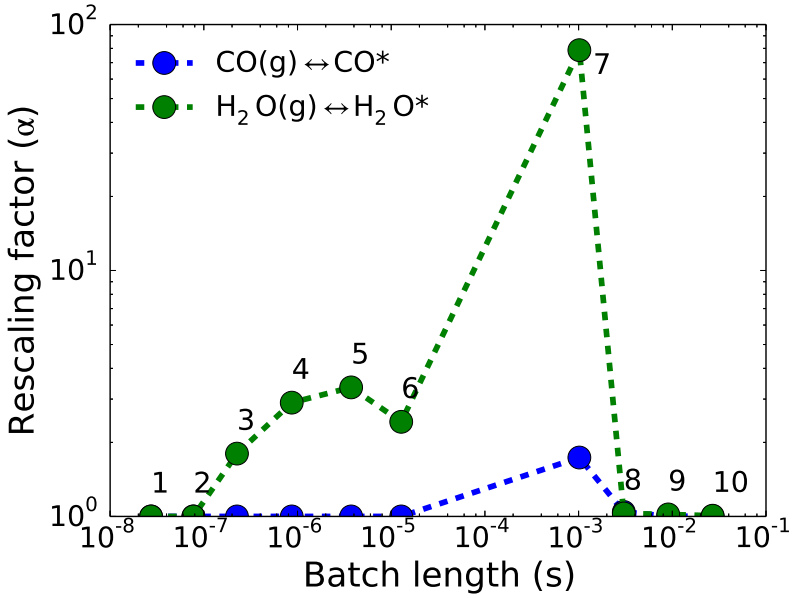


Figure 3.9: Rate constant rescaling factors used in the WGS model.

that the rate determining steps are correctly identified.

For the purposes of uncertainty quantification, the RDS should be investigated in more detail. Specifically, experimental evidence in the literature suggests that H_2O^* dissociation may occur on the support material rather than the metal[38, 99, 191]. Refinement of the model hierarchically[183] could revise the entire water-gas shift mechanism to include the support and possibly the metal/support interface. To this end, progress in modeling the metal support interface could be leveraged[10, 26, 135]. These topics are outside the scope of this computationally-focused paper.

3.4.3 Quantification of acceleration

To quantify the acceleration of the three strategies, we define speedup factors for parallel processing (S_{pp}), rate constant rescaling (S_{rcr}), and LRSA (S_{LRSA}). To decouple the effects of each strategy, we define each speedup factor as the ratio of the CPU time required for the nominal simulation method and the CPU time required if only that one strategy were used. The nominal simulation method, in which all three strategies are absent, would be running one trajectory on a single processor with no

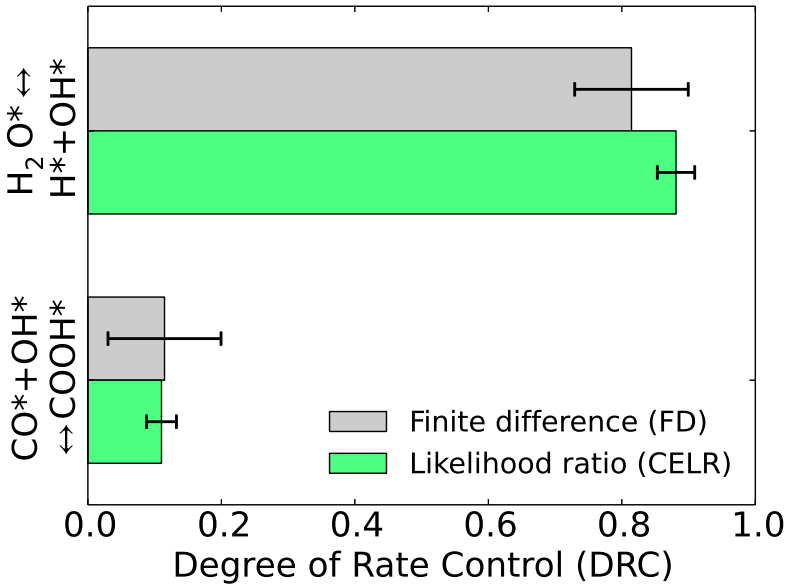


Figure 3.10: Sensitivity analysis data for the WGS model. Only the slow reactions with non-negligible sensitivities are shown. Likelihood ratio estimates are benchmarked against the finite difference (FD) estimates.

rate constant rescaling using finite difference sensitivity analysis. In practice, synergies between the strategies would occur.

For parallel processing, the length of the cumulative trajectory (t_{f}) is compared to the length of a single trajectory that would be required to achieve the same level of sampling, i.e., the sum of the lengths of all steady state batches. Equation 3.10 relates the length of trajectories to the level of sampling (N_{batches}) and the number of trajectories (n_{traj}). S_{pp} is always less than n_{traj} because for each trajectory, the first batch of every trajectory is discarded from the data. In the extreme limit of a large number of processors, the speedup is limited by the requirement that $N_{\mathrm{bpt}} \geq 3$. Given our definitions, Equation 3.10 is reformulated as

$$\mathrm{S}_{\mathrm{pp}} = \frac{N_{\mathrm{batches}} + 1}{N_{\mathrm{bpt}}} \leq n_{\mathrm{traj}}. \tag{3.29}$$

For rate constant rescaling, the CPU requirements of simulating the trajectories are compared to what they would be if the trajectories were simulated without rescaling. Reaction rate rescaling has little effect on the CPU time per KMC event, but it

significantly reduces the number of events needed to advance a given time interval. We compute the speedup factor from simulation data as

$$\mathrm{S_{rcr}} = \frac{\sum_{i=1}^{n_{\mathrm{iters}}} t_{\mathrm{f}}^{i} \left(\frac{\mathrm{CPU}^{1}}{t_{\mathrm{f}}^{1}} \right)}{\sum_{i=1}^{n_{\mathrm{iters}}} \mathrm{CPU}^{i}} \leq \max_{i} \alpha_{i} \ . \tag{3.30}$$

CPU^{i} is the CPU time required for iteration i. The numerator extrapolates the CPU requirements of the first iteration, where no rate constant rescaling has been implemented, to the other iterations. The denominator quantifies the CPU requirements of the simulation with rescaling. We also note in Equation 3.30 that $\mathrm{S_{rcr}}$ is limited to the degree by which rescaling of the rate constants occurs, and therefore, it heavily depends on the level of stiffness in the original system.

We define the speedup factor for LRSA to be the ratio of the total simulation times required for a hypothetical finite difference scheme and the likelihood ratio method. Central finite difference sensitivity analysis requires the nominal run as well as two additional runs for each of n_{slow} slow reactions. We do not perform finite difference on fast, equilibrated reactions because they are known to be insensitive. The simulation length required to achieve a desired level of confidence in the sensitivity estimates depends on the variance of the sensitivity method. For a fixed number of trajectories and batch length, the variance the finite difference and LRSA methods scale inversely with simulation time. Whereas previous formulations of LRSA using ergodic averaging showed variance that remained constant with simulation time[8, 82], our use of batches and steady state averaging[230] in Equation 3.27 achieves an inverse relation. We define VR as the ratio of the variance of LRSA to that of finite difference. We expect LRSA to have low variance for the slow reactions, as observed for our water-gas shift network. Arampatzis et al.[8] conducted extensive benchmarking of the variances of different sensitivity analysis methods. Based on their findings at short simulation times, we choose $\mathrm{VR} = 10$ to be an appropriate value.

The speedup factor ($\mathrm{S_{LRSA}}$) is then

$$\mathrm{S_{LRSA}} = \frac{(1 + 2n_{\mathrm{slow}})}{\mathrm{VR}}. \tag{3.31}$$

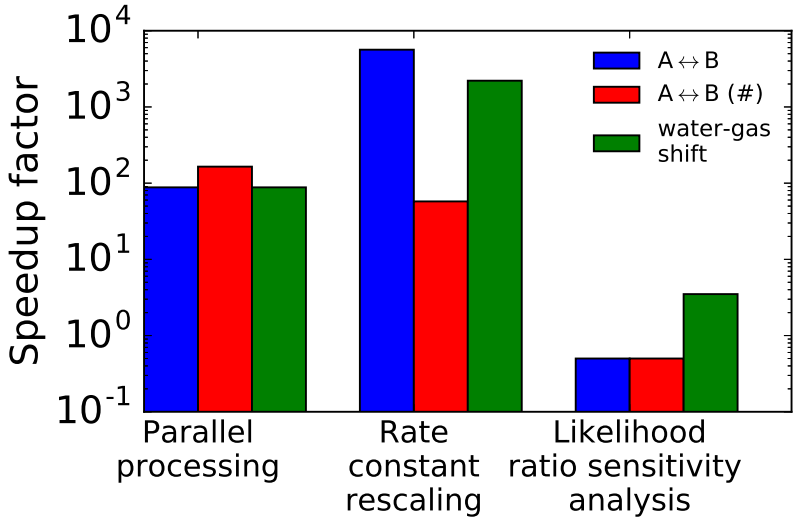


Figure 3.11: CPU speedup factors for the A→B, A→B (#), and WGS models, as defined in Section 3.4.3.

The functional form of Equation 3.31 suggests that if VR > 3, then $S_{\text{LRSA}} < 1$ when n_{slow} is small and $S_{\text{LRSA}} > 1$ when n_{slow} is large. Therefore, the sensitivity analysis method should be chosen depending on the size of the reaction network. For the A→B and WGS models, the speedup factors are shown in Figure 3.11. To emphasize that the speedup factors depend on the number of trajectories used as well as the initial level of stiffness, the A→B model is run with different parameters than before ($n_{\text{traj}} = 192$ and $k_1 = 10^3$ atm$^{-1}s^{-1}$; model denoted as (#)). For both the A→B and WGS models, rate constant rescaling has the largest speedup due to the severe stiffness of the original system. For the A→B (#) model, the increase in the number of trajectories combined with the reduced stiffness in the initial system causes the speedup factor for parallel processing to become the largest. The speedup factor LRSA is the smallest for all three models. Given our assumptions, LRSA is worthwhile for the WGS system, but not for the A→B model, as the water-gas shift reaction contains a greater number of slow reactions.

3.5 Conclusions

A three-pronged strategy for accelerating KMC simulation, while obtaining reliable statistics, was developed and implemented in a Python interface for Zacros, the commercial KMC software package. Steady state is achieved through an iterative process, in which the trajectory length is increased in successive iterations and convergence criteria are checked after each iteration. A batch means strategy is used to ensure that the variance of the estimate of the rate is sufficiently low and that adjacent batches are uncorrelated. Our specific strategy allows for multiple trajectories to be simulated in parallel, thus decreasing the sampling needed from a single trajectory and reducing real clock computational time. Rate constants are rescaled in successive iterations with error control to drastically accelerate simulation. After convergence is achieved, the batch length is identified as the decorrelation time, a key numerical parameter that informs the likelihood ratio sensitivity analysis method for identifying the rate-determining step. Additionally, the reaction rate rescaling procedure identifies partially equilibrated reactions for which the sensitivities are set to zero.

We tested our approach on a model system and a water-gas shift model on Pt(111), for which the reaction rates exhibit small and large variances, respectively. We find that simultaneously requiring a small confidence interval on the rate and negligible autocorrelation between batches are necessary to reliably guarantee convergence. Due to high variance in real chemistries, such as the water-gas shift reaction, large batch lengths are needed to manage the variance in the estimate of the mean of the rate. The CELR sensitivity analysis is also needed for practical application of the WGS model due to high variance in the rate.

All three acceleration strategies provide significant speedup, although the degree and relative importance depends on the number of trajectories (and thus processors) used as well as the parameters in the model. Savings of several orders of magnitude are typical for rate constant rescaling and parallel processing. Likelihood ratio sensitivity analysis achieves modest speedup relative to what would have been required with finite difference for the water-gas shift system, but not for the much smaller A $\rightarrow$ B

model, as this speedup scales linearly with the reaction network size. Given the high computational cost of KMC simulation, these strategies are essential for conducting real chemistry studies.

Chapter 4

OPTIMIZATION OF TRANSITION METAL CATALYST FACET STRUCTURE: APPLICATION TO THE OXYGEN REDUCTION REACTION

4.1 Abstract

Predicting the optimal catalyst structure has been a long-standing goal but typically, an active site on a uniform surface is a priori assumed. This assumption fails for structure sensitive chemistries, such as the oxygen reduction reaction (ORR). Here we develop an approach to predict the optimal catalyst structure by identifying the active site and the density and spatial arrangement of such sites while minimizing the surface energy. Peak performance is unattainable due to lack of suitable active sites on low index planes, as well as geometric and stability constraints. A random array of vacancies results in modest performance enhancement compared to ideal facets; in contrast, engineering defect sites of suitable geometric motif with maximum density in disordered structures significantly increases catalyst performance. We apply this methodology to the ORR on defected Pt(111), Pt(100), Au(111), and Au(100) surfaces and provide insights into structure optimization.

4.2 Introduction

Highly dispersed metal nanoparticles are used ubiquitously as catalysts. Substantial effort has been devoted over the past two decades to determine how reactions are affected by the size and shape of nanoparticles and how best to maximize catalyst activity[37, 111, 223]. Techniques for synthesis of shape selecting nanoparticles have been developed[178, 196] and applied to various chemistries including formic acid oxidation[235], acetylene hydrogenation[106], NO reduction[173], benzene

hydrogenation[22], trans-to-cis isomerization[117], and 2-propenol oxidation[144] to mention a few. Despite these advances, engineering the nanoparticle size and shape from first principles has remained elusive. In recent, exciting studies on ORR on Pt(111)[28] and PtNi/C nanoparticles[50], it was shown that surface defects can enhance activity. Similarly, a multiscale computational study demonstrated that patched bimetallic surfaces with vacancies on the top layer can deliver unprecedented boost in activity for the ammonia decomposition reaction[75]; size and shape of defects were obviously important but how to predict optimal structures remains elusive. These early studies underscore a rather novel and unexplored means of enhancing catalyst activity whereby one purposely engineers defects on catalyst surfaces. Arbitrary defects, by contrast, may lead to a decrease in activity[30].

The vast structure-activity design space, where surfaces are purposely engineered with atomistic detail and their catalytic properties are computed on the fly, imposes a major computational challenge but promises to unlock unexplored structures with improved performance. Due to the computational cost of first-principles methods, such as density functional theory (DFT), direct evaluation of the activity of numerous structures is beyond current supercomputer power. Instead, a surrogate microstructure-energy model is necessary. A recent breakthrough in correlating the binding energy with microstructure entails the work of Sautet and co-workers who related the binding energies of ORR intermediates to the generalized coordination number (GCN) of a binding site[29, 28, 31]. An earlier study describing adsorption of CO on Au clusters with the coordination number of the binding site (a geometric descriptor) provides further evidence that activity and microgeometry can be correlated[145]. These structure-scaling relations enable high throughput screening of microstructures by computing GCNs simply through counting of first and second nearest neighbors. By constructing volcano curves vs. the GCN, Sautet and co-workers showed that an optimal site on Pt(111) crystal exists and demonstrated experimentally that defects indeed increase the catalyst activity, compared to the ideal surface. However, the experimental

activity enhancement was modest compared to a structural model consisting of identical, single active sites corresponding to the volcano peak. This raises fundamental questions of the predictive ability of models and whether microstructures consisting of closely packed optimal active sites were experimentally made and stable, given that highly defected structures should possess higher surface free energies than ideal low Miller index planes. Despite these exciting developments, the maximum density of defects one can geometrically arrange in a structure, the effect of spatial organization of defects, and the maximum activity achieved in an experiment remain unclear. Addressing these questions requires further conceptual advances beyond the structure-based volcano curve. Hanselman et al.[78] proposed active sites given a descriptor but did not consider stability requirements, which are critical, given that previous studies have related active site stability to activity, and in some cases, found a tradeoff[179, 240].

Here, we introduce a framework that can predict the microstructure resulting in maximum activity of a catalyst surface along with its stability using a multi-objective optimization approach. To the best of our knowledge, this is the first study to design atomistic structure dynamically in terms of both optimal site-activity and surface stability and address fundamental questions that may close the gap of models and experiments. We demonstrate the approach on Pt(111), Pt(100), Au(111), and Au(100) facets for the ORR. ORR at the cathode of proton-exchange membrane (PEM) fuel cells is an important structure sensitive chemistry[70, 118, 131, 162]. Currently, widespread commercialization using Pt/C catalysts is hindered because of slow kinetics (a high overpotential), the high cost of Pt, and catalyst degradation under harsh acidic conditions[58, 98, 226]. Consequently, there remains ample room for catalyst design to advance the PEM fuel cell technology[208, 236]. We report a rather unexpected result: the optimal active site changes among metals and facets of a single metal and with the density of active sites but only a small number of geometrically distinct sites contributes to activity. Disordered structures of such (highly packed) active sites deliver best performance. Expectedly, creation of defects increases the surface free energy. However, numerous metastable structures, which possess lower activity than that of

the peak of the volcano curve, exist. Combined, the lack of sites on low index planes that match the peak of the volcano curve, packing constraints, and surface stability rationalize that (geometric, volcano-based) peak activity is unattainable. Importantly, metastable structures are still an order of magnitude (or more) superior to ideal facets. In contrast, vacancies in the top layer at the same density as those of our optimal structures provide only modest performance enhancement. Making the right geometric motif of defects and close packing such defects are essential and key to materializing the benefits of defect engineering of crystal surfaces and provide a potential explanation of the modest activity seen experimentally.

4.3 Methods

The ORR mechanism is modeled as a series of four sequential elementary steps, each of which reduces an oxygen species, such as *OH and *OOH, with the addition of an electron and hydrogen ion. It is well known that the binding energies of the *OH and *OOH intermediates vary among facets[122, 215, 216]. Based on the structure-scaling relations, the activity of a site depends on its coordination environment. Section C.1 describes how a site's activity is computed based on structure-scaling relations using the GCN and experimental data. Structure dependent volcano plots for Pt and Au vs. the GCN are shown in Figure 4.1. *OH removal is rate determining on the left leg of the volcano while *OOH formation is rate determining on the right leg. Au is a weak binding metal and has an optimal GCN of 5.75, which is smaller compared to the GCN of the (100) and (111) planes. In contrast, Pt binds oxygen species strongly and its activity peaks at a GCN of 8.29, which is higher than the value on either perfect facet. Clearly, ideal facets are suboptimal and creating defects can be beneficial. Figure 4.1 indicates that potentially different strategies of crystal imperfection may be necessary for maximizing activity of weak and strongly binding metals that lie on either side of the volcano curve. We elaborate on this point below.

The structure of defected Pt(111), Pt(100), Au(111), and Au(100) facets, which maximizes current density while minimizing surface energy, is predicted using a two

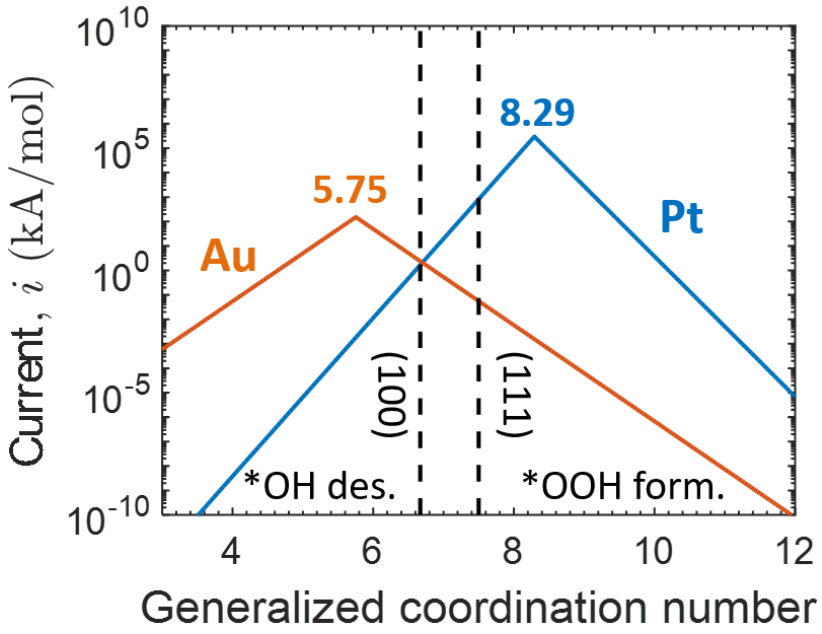


Figure 4.1: Volcano map for ORR activities on Pt and Au. Current i is as defined in Equation 4.1. The GCN of the optimal site is 5.75 and 8.29 for Au and Pt, respectively. The vertical dashed lines at GCN = 6.67 and GCN = 7.5 show the GCNs of the (111) and (100) planes for reference. Clearly, ideal facets are suboptimal; for Au and Pt one needs under and over coordinated sites to enhance activity.

level optimization procedure. First, simultaneous optimization of activity and stability is carried out (by weighing activity vs. stability through a weight, ω), as detailed in the Methods section. Following this optimization, the surface is quenched to a local energy minimum on the coarse-grained potential energy surface.

4.3.1 Coordination numbers as a descriptor of activity and stability

We compute the current i of a site as

$$i = i_c * \exp\left(-\frac{\Delta G_{\text{ORR}}}{k_B T}\right) \tag{4.1}$$

The constant $i_c = 3.68 * 10^{11}$ kA/mol is used to match the experimental value of the Pt(111) current density at a cell potential of $U = 0.9\ V$ and temperature $T = 298\ K$ reported by Calle-Vallejo et al[28]. The current density j of a microstructure is then computed from the activities of all surface sites normalized by the exposed surface area A_{surf}.

$$j = \frac{1}{A_{\text{surf}}} \sum_k i_k \tag{4.2}$$

Equations C.5 and C.6 in conjunction with Equation 4.1 describe the structure sensitivity of current density.

The stability of an extended surface is approximated by its surface energy (γ), which is computed by dividing the formation energy (E_{form}) of a slab by the exposed surface area.

$$\gamma = \frac{E_{\text{form}}}{A_{\text{surf}}} \tag{4.3}$$

Formation energy is rapidly calculated using a surrogate model derived from tight-binding theory (TBT)[53, 214].

$$E_{\text{form}} = E_{\text{coh}} \sum_k \left(1 - \sqrt{\frac{\text{CN}_k}{\text{CN}_{\text{max}}}}\right) \tag{4.4}$$

The summation is taken over all atoms k, each with coordination number CN_i. $\text{CN}_{\text{max}} = 12$ is the maximum coordination number in an fcc lattice. E_{coh} is the cohesive energy of the metal. In order to validate Equation 4.4, DFT calculations are performed as detailed in C.2.

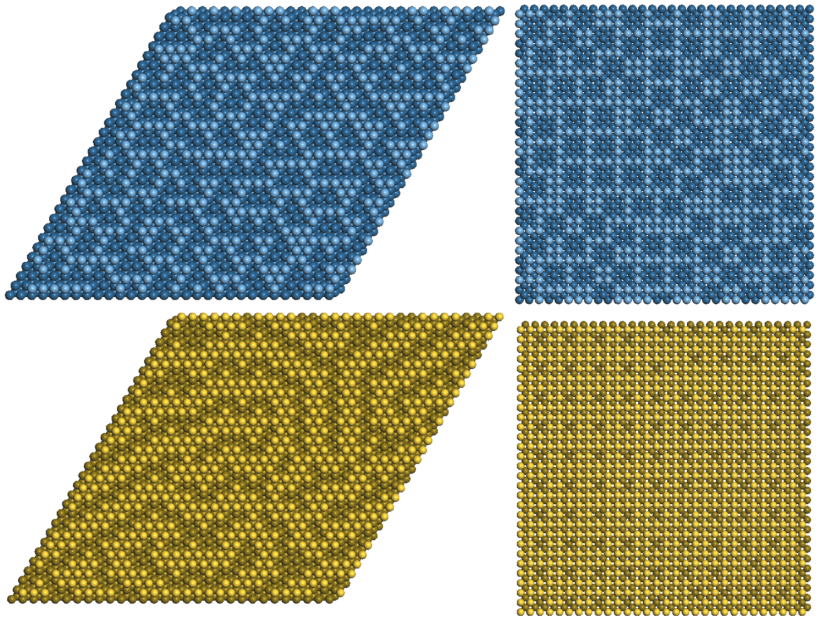


Figure 4.2: **Strcutures of defected crystals.** $p(30\times30)$ structure of Pt(111) (top left), Pt(100) (top right), Au(111) (bottom left) and Au(100) (right) slabs with defects in the top layer 1 (light colored atoms). Atoms in layer 1 are added and removed during optimization. Atoms in layers 2, 3, and 4 are held fixed. The structures shown were the highest activity structures for each surface achieved after the quenching step.

4.3.2 Catalyst structure and optimization

The catalyst surface is modeled as a four-layer (111) or (100) slab, shown in Figure 4.2, of a $p(30\times30)$ unit cell size and with periodicity in the x-y directions. The layers in the slab are indexed from top to bottom. The top layer 1 may include atoms or vacancies, i.e., it is a defected layer, whereas deeper layers contain no defects (vacancies in deeper layers can easily be considered in our model but their geometric accessibility should also be taken into account). Each microstructure is uniquely described by a column vector $\vec{c}$ which lists the occupancies (0 if metal atom is absent or 1 if present) of the metal atoms at each lattice position. Each exposed binding site contributes to the overall activity dependent on the atom's GCN.

For each microstructure, the GCNs for all atoms in the slab are computed rapidly using graph theory. Furthermore, we exploit the sparsity of the adjacency matrix describing the nearest neighbors to significantly accelerate computations. Details are

described in C.3. Given the GCN of a site, the binding energies of OH* and OOH* are computed according to Equations C.5 and C.6, respectively, and are used to estimate the site specific activity (Equation 4.1). The activity of the structure is computed by summing over each active site (Equation 4.2). The structure's surface energy and total activity are non-dimensionalized to $\mathcal{E}$ and $\mathcal{I}$, respectively for numerical convenience, as detailed in C.3.

We perform a two level optimization. First, we simultaneously minimize the surface energy and maximize activity using a weighted sum for the objective function, as is standard in multiobjective optimization[46, 132].

$$\mathcal{F} = \omega * \ \mathcal{I} + (1 - \omega) * \mathcal{E} \tag{4.5}$$

$0 \leq \omega \leq 1$ is a "weighting" factor that emphasizes surface energy minimization ($\omega = 0$) or activity maximization ($\omega = 1$). A simulated annealing Metropolis Monte Carlo algorithm is applied and at each event, a position in the top layer of the slab is chosen, and a move, which changes the occupancy, is attempted (an occupied site becomes empty and vice versa) according to the Metropolis algorithm using the change in the objective function $\Delta\mathcal{F}$ (see details in C.4). Inverting the sign for activity in the objective function causes our algorithm to create structures with activity lower than that of an ideal surface, demonstrating how specific types of defects, rather than arbitrary ones, are needed to achieve higher activity. An example is shown in Section C.7.

Given that the multi-objective optimization strikes a balance between competing factors, the resulting structures can be unstable to small perturbations, especially when a large weight on activity is placed. For this reason, following the multi-objective optimization, relaxation is performed to locally minimize the surface energy whereby atoms in the top layer are allowed to diffuse using a downhill energy minimization approach. The structure converges to the nearest local minimum on the coarse-grained potential energy surface, typically, at the expense of some loss in activity. Similar to the multi-objective optimization, we take advantage of the sparsity of the adjacency

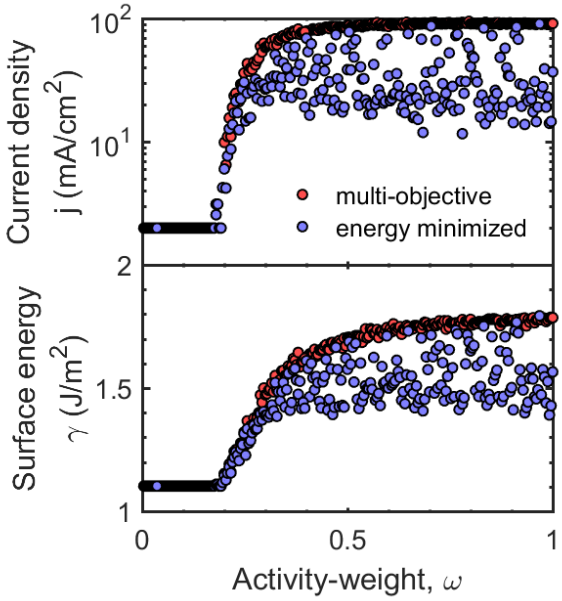


Figure 4.3: Current density and surface energy of numerous defected Pt(111)-based crystals containing different active sites and/or density of active sites. Current densities (Equation 4.2) and surface energies (Equation 4.3) from optimization are plotted against the activity-weight ω. Red points show values following the multi-objective optimization. Blue points show data after the energy-only minimization is also applied.

matrix to speed up the algorithm (see C.3). The entire algorithm is summarized in a pseudocode in C.4.

For a given metal-facet, simulations are run for ω values using a random initial structure and random seed for each one. The inherent randomness of the optimization results in some structures having superior properties to others (see fluctuations in blue points in Figure 4.3). Structures of high current density and low surface energy are most interesting. C.5 describes how a Pareto optimality criterion is used to select these structures.

4.4 Results

4.4.1 Active Site and Optimal Structures

Surface energies and current densities for Pt(111)-based metastable structures resulting from the first optimization are shown in Figure 4.3 (red points). The other surfaces follow similar trends and are included in the supplemental information (Figure

C.7). A nearly continuous set of solutions between the ideal facet ($\omega = 0$) and the most active (but not necessarily stable; $\omega = 1$) structures is obtained. For ω ($>\sim 0.4$), the activity and surface energy vary only slightly with ω; primarily activity is optimized. In contrast, for low values of ω ($<\sim 0.2$), the stability of the structure dominates optimization. The transition between the two regimes is fairly abrupt. Subsequent quenching to a local minimum energy lowers the surface energy at the expense of activity, typically by a factor of 50-75% (Figure 4.3; blue points). Diffusion of surface atoms eliminates some active sites created during the multiobjective optimization. Error bars on the activities in Figure 4.6 quantify the variability in how much activity is lost by destroying active sites.

The results indicate a profound effect of introducing defects to a structure whereby the catalyst activity can increase by nearly two orders of magnitude compared to the stable, low index plane. At the same time, they indicate that stability and details of the microstructure matter. We elaborate on these topics below.

The highest activity structures following the quenching step are shown in Figure 4.2. Analysis of the structures identifies the active sites, shown in Figure 4.4. For Pt(111), several types of cavities are active. Figure 4.4a shows a 5-atom vacancy which exposes two adjacent active sites with GCN of 8.25, which is near the Pt peak of 8.29. Structures at intermediate values of ω have this cavity as the predominant active site. A 4-atom vacancy, shown in Figure 4.4b, has a slightly higher GCN of 8.33 and is more active, but contributes less to overall activity because there is only one active site in the cavity, i.e., due to a lower density of such defects. Structures at low ω consist of disjoint cavities (Figures Figure 4.4a and Figure 4.4b). At higher activities, denser packing of active sites is required, resulting in a predominant cavity shown in Figure 4.4c. It includes vacancies, which branch out into other cavities, allowing for a dense network of active sites. The most active structures consist of a disordered ensemble of these active sites, as can be seen in Figure 4.2.

Calle-Vallejo et al. proposed that the active site on defected Pt(111) was either a six-atom or five-atom vacancy exposing a second-layer atom with a GCN of 8.00 or

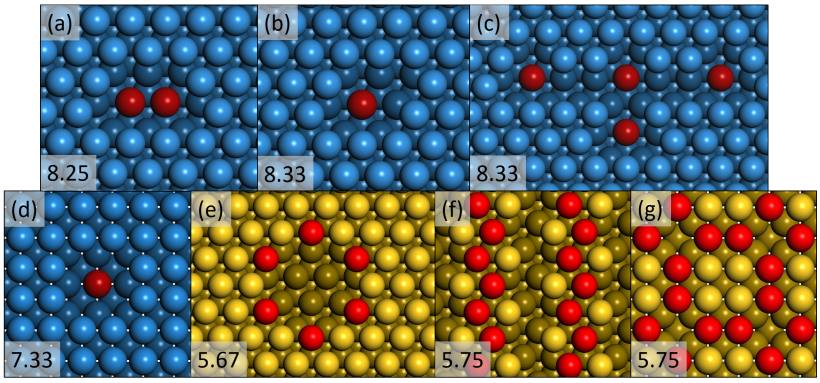


Figure 4.4: **Top-down views of the active sites for each surface**. Only the top and next layers, when its site is exposed, are shown. (a)-(c) are Pt(111) sites. (d) is a Pt(100) site. (e) and (f) are Au(111) sites. (g) shows Au(100) sites. Red atoms indicate the most active sites, which occur in the second layer for Pt and the top layer for Au. Blue atoms are Pt and gold atoms are Au. Darker colored atoms indicate atoms in the bottom layers. Numbers in the bottom left of each panel indicate the generalized coordination number (GCN) of the active site. Each surface has an optimal GCN that differs from the ideal one (peak of volcano) and is determined from symmetry and the metal itself. Maximum activity requires random packing of sites to maximize the number of sites per unit area and the optimal active site may change with increasing packing of sites and thus with catalyst activity.

8.17, respectively. The 5-atom vacancy is similar to the structure in Figure 4.4a and can be packed similarly on the surface. A maximal ordered packing of this site would result in activity 14 times higher than Pt(111). By comparison, our best structure is 46 times more active than Pt(111) (see below). We attribute the better performance of our best structure to the slightly higher (about a factor of 2) activity of our optimal active sites and the disordered packing of various types of active sites. Our exhaustive and systematic approach predicts more active sites than intuitive choices (i.e., closer to the volcano peak) that can be packed densely on the surface. Additionally, our approach is general to other metals and structures.

The Pt(100)-based surfaces show more order. A four-atom vacancy exposing a single active site in the second layer is the only active site found (Figure 4.4d). The differing activities of Pt(100) surfaces are attributed only to the differing densities of this active site. The GCN of the active site is only 7.33, which falls short of the Pt peak. The (100) geometry is unable to possess a site with a higher GCN by introducing vacancies in the top layer.

Similar to Pt(111), Au(111) has multiple types of active sites. For the less active structures, large cavities shown in Figure 4.4e are present. The active sites reside on the top layer edge sites along the cavity. All atoms have coordination numbers of 7 or higher and therefore, surface stability is relatively maintained. The most active structures contain a more active but less stable configuration of sites, shown in Figure 4.4f. Channels of vacancies expose edge sites on the top layer in high density. Defected Au(100) has a periodic pattern of vacancies shown in Figure 4.4g. Because the optimal GCN is not very different from the ideal (100) surface, only a few vacancies are needed. For each of the active sites shown in Figure 4.4, the predictive accuracy of our model is assessed with DFT, as detailed in C.2 of the supplement. Validation of the active sites allows for model refinement in an efficient hierarchical refinement approach[184, 210, 220].

Current density and surface energy are plotted against each other for the Pareto optimal points in Figure 4.5. These structures span from low activity/low-energy

structures to highly active/high energy structures. All Pareto fronts have linear portions, which are indicative of a single type of active site being added in higher density as activity increases. The constant slope indicates proportional increases in surface energy and activity. The Pt(100) and Au(100) curves are entirely linear, indicating that only one type of active site is present, consistent with structural analysis presented above (Figure 4.4). In contrast, the Pt(111) and Au(111) Pareto fronts have kinks where the active site being added to the surface changes with increasing activity, consistent with the structural analysis above. The Pareto front for Au(100) has a large gap. Only two structures are relevant: the ideal facet as well as the periodic structure shown in Figure 4.4g. Periodicity does not allow intermediate structures with lower density of active sites. Overall, only a relatively small number of distinct geometric motifs exists on each metal facet. Clearly, creation of defects giving very active structures drives up significantly the surface energy (Figure 4.5). Yet, Figure 4.3 indicates that even though the optimal structures possess higher surface energy, numerous metastable structures of lower surface energy exhibit a nearly order of magnitude increase in activity (blue points at large ω). This offers promise that profound performance improvement is feasible via defect-engineering of experimental structures.

4.4.2 Catalyst Activity and Stability

The increase in current density due to defects is summarized in Figure 4.6. For Pt(111), our most active structure consists of 49% vacancies on the top layer and is 46 times more active than an ideal Pt(111) surface. This is about an order of magnitude less active compared to the peak of the volcano curve due to the inability to create such an active site on Pt(111) and geometric constraints, i.e., inability of having each site being the most active. By comparison, Calle-Vallejo et al. observed experimentally a 3.5-fold increase in activity due to defects[31]. Our best structure is about an order of magnitude more active than the experimentally measured one. We attribute this difference either to an experimentally lower density of defects or

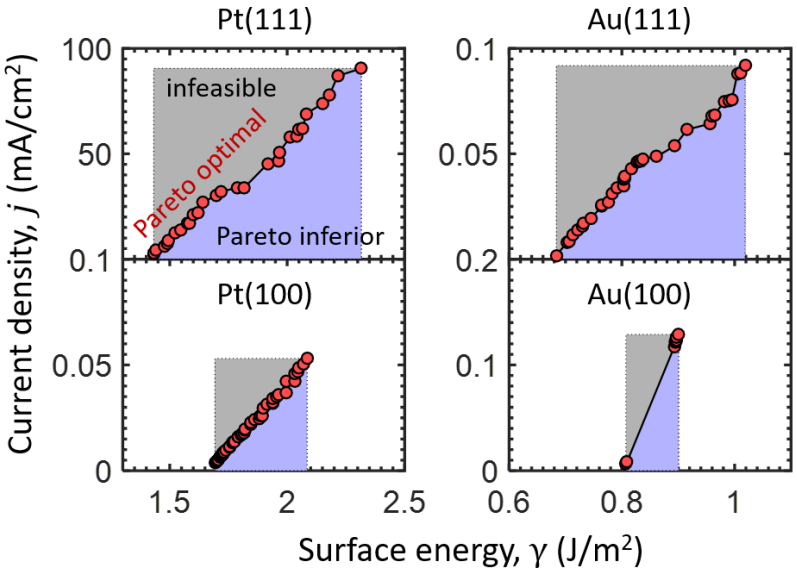


Figure 4.5: Current density j (Equation 4.2) vs. surface energy γ (Equation 4.3) of defected surfaces. The black lines trace the data of the optimal structures indicating the Pareto tradeoff between stability and activity. Straight lines indicate that activity increases due to an increase in the density of the same active site; kinks in linearity are indicative of creation of more than one active site. The most active structures have the highest surface free energy and are thus less stable.

defects of lower activity; we cannot infer the precise reason without detailed microscopic characterization of the experimental structures. However, we provide further insights by creating a random structure of defected Pt(111) with the same number of (49%) vacancies in the top layer as in our optimal structure. Then we subject the structure to annealing (by gradually decreasing the temperature in a Monte Carlo simulation) with constant Pt loading of adatoms (Figure C.5). This situation could relatively easily be realized in a real experiment, where a certain number of adatoms is added on Pt(111), for example via atomic layer deposition (ALD), and the structure is then slightly annealed. The resulting average activity of the annealed structure is surprisingly comparable (slightly lower than) to the experimental value and double of that of the perfect Pt(111) (Figure C.5). Our results underscore that the total density of vacancies or adatoms on the top layer is an important but not sufficient factor for improving activity. Furthermore, we hypothesize that the experiments had probably a suboptimal distribution of defects. Microscopic engineering of the right microgeometry of active sites and maximizing the density of such defects are crucial to materialize fully the benefit of crystal imperfection. Randomly created structures with the correct density of vacancies or adatoms give only a modest boost in activity.

The other surfaces experience differing degrees of improvements when defects are introduced. Defects in Pt(100) increase the activity by up to a factor of 15, which is threefold less than the enhancement on Pt(111) because the active site does not reach the volcano peak and therefore gains in current density are modest. Au(111)'s activity can be enhanced by almost 3 orders of magnitude due to how densely the active sites are packed along edges in the optimal structure. Au(100)'s activity, by contrast, benefits less from defects because it is already reasonably active relative to the volcano peak. It is clear that the higher symmetry of the 111 plane enables a higher degree of engineering, by introducing defects, and activity-improvement, compared to the 100 facet. For Pt, even the best structures fall short (by an order of magnitude) of reaching the peak activity of this catalyst. Pt cuboctahedral particles with surface defects are highly desirable. In contrast, defected cubic particles of Au approach closely

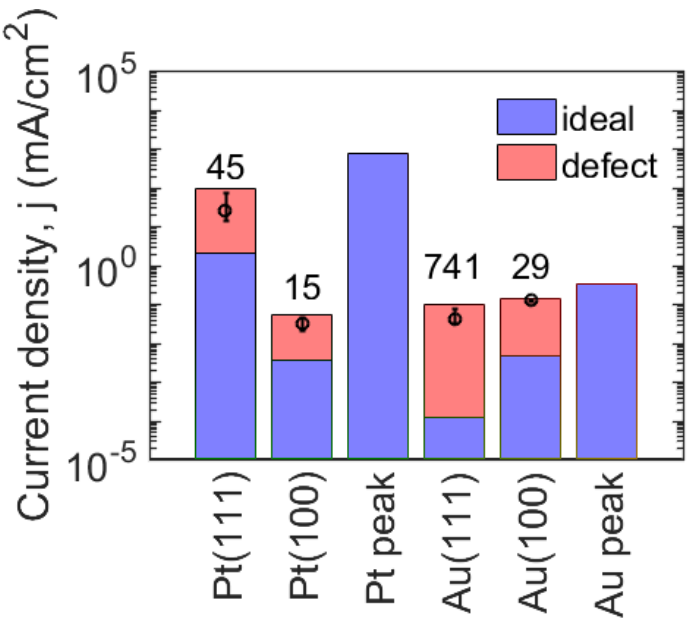


Figure 4.6: **Activity of defected, metastable structures compared to ideal crystals**. Numbers above the bars indicate the ratio of activities between the perfect, non-defected (blue bars) and most active, defected metastable surfaces (red bars). Circles and error bars are 90% confidence intervals around the median of activities of annealed structures (average of blue points in the high activity regime of 4.3 and Figure C.4), which are more realistic than the top activity of the bars. As a reference, Pt and Au peaks correspond to the optimal site of the volcano curve assuming a density of such sites to be that of the (111) facet; such densities are not geometrically possible. Defect engineering can profoundly increase activity over ideal crystal facets but site geometry, density, stability should carefully be considered.

the optimal performance of this metal. Yet, Au is significantly less active than Pt even with introduction of defects.

The proposed approach can be combined with "activity maps"[79, 150, 183] to predict the structure of bimetallic catalysts, and can easily be extended to multiple low Miller index planes of nanoparticles[55, 56, 85, 96] or entire nanoparticles via the Wulff construction[15, 215, 216] and eventually discrete descriptions of the surface modeled with kinetic Monte Carlo simulations[75, 123, 204, 237] and selectivity problems[45].

4.5 Conclusions

We have introduced a framework to enable atomistic level materials engineering. The framework predicts optimal catalyst performance by simultaneously identifying both the optimal active site of a surface facet of a catalyst and the geometric arrangement of such sites that maximizes their density, while accounting for catalyst stability. Our approach offers insights unobtainable by intuition alone. We have applied this framework to predict optimal defects on Pt(111), Pt(100), Au(111), and Au(100) surfaces in catalyzing the ORR chemistry of significance to fuel cells.

The results are rather revealing. Each material and facet has a different optimal active site whose geometric structure is not intuitive. This is unlike volcano curves constructed in the past assuming the same active site for all materials. For materials on the left of the structure-based volcano curve (activity vs. generalized coordination number), over coordinated sites (vacancies) are needed, whereas on the right of the volcano curve, under coordinated sites along edges are needed. The active site on Pt(111) and Au(111) changes with increasing defect density due to packing constraints. Interestingly, optimal facet structures contain a disordered ensemble of a small number of distinct sites. Our results indicate that optimal structures are significantly inferior to the peak of the volcano due to lack of ideal active sites, packing constraints on low index planes and higher surface energy compared to the ideal facets. Yet, metastable, defected structures can still be significantly better than perfect crystals. In contrast, randomly put structures give marginal activity improvements. Defect engineering,

whereby active sites of the right geometric motif are created and stabilized, is thus imperative to materialize significant performance enhancements.

For the studied ORR catalysts, the activity is ranked Pt(100) < Au(100) < Pt(111) < Au(111). Geometrically, all surfaces other than Pt(100) are able to produce nearly optimal sites but the density of these sites limits activity. The activity of Au improves much more than Pt because the low-coordinated active sites on Au are easier to produce in abundance than the high-coordinated active sites on Pt. Similarly, the (111) facets exhibit higher degree of engineering than the (100) facets because the hexagonal symmetry offers greater flexibility compared to the square symmetry.

The proposed structure optimization allows for the first time identification of the active site and of optimal arrangement of such sites and comparison of different materials and facets. This capability can be extended to any chemistry to enable significant improvements of catalysts.

Chapter 5

AN ONLINE MACHINE LEARNING APPROACH TO OPTIMIZING ACTIVE SITE COUPLING

5.1 Abstract

Computational prediction of optimal catalyst structure is an important goal in catalyst design. A major challenge is that active sites on the catalyst surface are coupled due to complex phenomena such as bifunctionality, adsorbate interactions, and diffusion. Modeling these effects requires computationally demanding kinetic Monte Carlo (KMC) simulation, making it infeasible to carry out optimization. To overcome this challenge, we develop a surrogate model using machine learning tools regressed to KMC data. The surrogate model captures the complex structural dependencies among sites typically modeled via KMC at comparably negligible computational cost. However, inadequacy in the training data used to regress the surrogate model may lead the optimization to predict false optima. Utilizing an online machine learning approach, we refine the surrogate model with additional data obtained by evaluating the predicted optima with KMC. The optimization is repeated and the surrogate model is refined iteratively. In doing so, the true global optimum is eventually identified and predicted accurately by the surrogate model. We apply our method to a prototype chemistry exhibiting bifunctionality similar to that of the ammonia decomposition reaction on defected NiPt. Catalyst activity is improved by a factor of 5 relative to a set of randomly generated structures, while the computational cost of optimization is reduced by 3 orders of magnitude relative to a brute force approach.

5.2 Introduction

Highly dispersed metal nanoparticle catalysts are inherently heterogeneous, consisting of different sites where adsorbates can bind and react. The various sites often exhibit dramatically different activities and their relative abundance changes with catalyst size and shape, resulting in structure sensitivity. This principle has been exploited to engineer catalyst activity and selectivity. Techniques for synthesis of shape selected nanoparticles have been developed[196, 178] and applied to various chemistries including formic acid oxidation[235], acetylene hydrogenation[106], NO reduction[173], benzene hydrogenation[22], trans-to-cis isomerization[117], and 2-propenol oxidation[144] to mention a few. Defects have been purposefully engineered for the oxygen reduction reaction [31]. Extensive work has been done to understand structure sensitivity computationally[223, 111, 37] with the eventual goals of optimizing catalyst structure and providing synthesis targets for experimental studies. This approach requires consideration of the identities of the active sites as well as their spatial arrangement.

The activity of a surface site is predicted by computing properties (e.g., binding energies) that affect reactivity. Density functional theory (DFT) is currently the best tool for this task, but its computational expense makes it a poor choice for exhaustive screening to identify the most promising active site geometries. Instead, DFT data can be used to build geometry-property relationships based on a site's coordination number[145, 156, 29, 28, 31, 126], Hamiltonians[74], or machine learning[221]. Brønsted-Evans-Polanyi (BEP) relationships[89, 151] extend binding energies predicted by screening methods to kinetic parameters for a reaction network of interest.

Current methods for screening active sites are limited to evaluating them one at a time, i.e., in isolation. In real systems, active sites are coupled through several phenomena, to which we collectively refer as active-site coupling. As shown in Chapter 4 and Hanselman et al.[78], the stability and geometric requirements of an active site affect the properties of nearby sites, excluding the possibility of every site having optimal properties. Bifunctional catalysts, in which different parts of a reaction mechanism occur on different sites[60, 157, 228, 136, 143], also challenge the existing

paradigm of evaluating active sites one at a time. Furthermore, diffusion[5] and adsorbate-adsorbate interactions couple the coverages and activities of nearby sites. Combined, these effects cause not only the density but also the spatial arrangement[49] of active sites to matter.

The effect of active-site coupling on catalyst activity can only be accurately evaluated using KMC simulation, whose application to structure sensitive chemistries is detailed in several review papers[183, 206, 200]. The NH_3 decomposition reaction is a quintessential example[101]. Although Ni and Pt are inactive for NH_3 decomposition, NiPt is the most active known single crystal. Previously, the high activity of NiPt was rationalized by well-established volcano plots[225] and the N binding energy on an ideal NiPt monolayer[79]. However, molecular dynamics[227], EXAFS[217], and other experimental methods[108] have shown that substantial defects are present in the NiPt structure. KMC simulations rationalized the high activity of NiPt based on an interplay between terrace and step sites[73, 75]. NH_3* dissociation on terrace sites and N* association on step sites are competing rate determining steps whose relative importance depends on reaction conditions. Lateral interactions were also shown to play a role[139, 219, 72].

Although the KMC studies mentioned above map catalyst structure to activity, there is currently no way to solve the inverse problem. An approach to optimizing activity with respect to catalyst structure in the presence of active-site coupling is still lacking because the computational cost of KMC simulation prohibits its direct use in optimization. Therefore, a surrogate model that captures active-site coupling at low computational cost is needed.

In this chapter, we develop a surrogate model consisting of a decision tree and a neural network regressed to KMC data. Initially, a database is built by randomly generating catalyst structures and evaluating their activity via KMC simulation. The surrogate model is trained on the structures in the database and subsequently employed in simulated annealing optimization to maximize catalyst activity. Due to inadequacies in the training set, the surrogate model is unlikely to evaluate the optimum

correctly. Therefore, an online machine learning approach is used wherein structures predicted by the optimization are evaluated with KMC and added to the database to refine the surrogate model. In doing so, computational effort is spent populating the database primarily with highly active structures, i.e., using importance sampling[125]. KMC simulation, surrogate model training, and catalyst structure optimization are performed in an iterative fashion. The iterative method converges when sufficient data has been added so that the surrogate model and KMC simulations agree for the simulated annealing optimum. Our approach is demonstrated for a prototype version of the ammonia decomposition reaction on defected NiPt catalysts.

5.3 Methods

In Section 5.3.1 we describe how catalyst structures are modeled and evaluated with KMC. Section 5.3.2 details the construction of the surrogate model and its regression to KMC data. Section 5.3.3 outlines the simulated annealing method used to optimize the catalyst structure. Section 5.3.4 combines these techniques into an online machine learning approach.

5.3.1 Kinetic Monte Carlo (KMC) simulation of defected monolayer catalysts

A prototype version of the NH_3 decomposition model of Guo and Vlachos[75] is chosen to demonstrate our methodology. A catalyst structure consists of a defected monolayer of Ni on a Pt substrate. The Atomic Simulation Environment (ASE)[86] is used to create a 4-layer fcc(111) slab consisting of 3 bottom layers of Pt and a top layer of Ni, periodic in two dimensions. The cell dimensions chosen for this study are $p(d \times d)$ where $d = 12$. Defects are introduced through vacancies in the Ni layer. The defected structure is represented numerically as a binary occupancy vector σ. Each of the $n = d^2 = 144$ Ni sites, indexed by $i \in \{1, 2, \cdots, n\}$, is occupied ($\sigma_i = 1$) if a Ni atom is present and unoccupied ($\sigma_i = 0$) if a vacancy is present instead. An example structure is shown in the left hand panel of Figure 5.1.

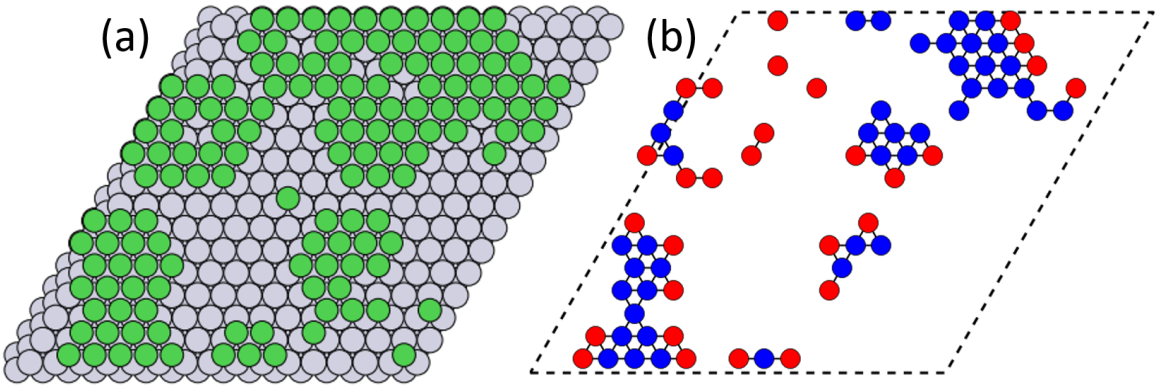


Figure 5.1: (a) An example catalyst structure. The occupancies of Ni (green) atoms on the Pt (grey) substrate are specified by a vector σ. (b) The corresponding KMC lattice with terrace (blue) and edge (red) sites.

The molecular structure is converted to a KMC lattice by identifying binding sites and their properties. The KMC lattice for the example structure in Figure 5.1 is shown in the right hand panel. Terrace and edge sites are identified in the Ni adlayer. Terrace sites are Ni top sites with 6 neighboring Ni atoms, whereas edge sites are Ni top sites with 4 neighboring Ni atoms and two neighboring vacancies adjacent to each other. All other Ni sites do not participate in the chemistry. Terrace and edge sites are given prespecified properties, as is the case when using DFT data[75]. For other chemistries, predetermined site types and properties would not necessarily be required if they can be computed on the fly based on a site's local environment[145, 29, 28, 31, 74, 221].

The elementary steps and rate constants of the reaction network are shown in Table 5.1. All rate constants are the same order of magnitude so that competing rate determining steps are present. Adsorption of A occurs on terrace sites whereas desorption of B occurs on edge sites. We emphasize that due to bifunctionality, the presence of a terrace or step site does not necessarily mean that it participates in the chemistry. Isolated terrace sites have no means of getting rid of adsorbed A except by desorption. Likewise, edge sites without nearby terrace sites have no source of B via diffusion.

Given a defected structure defined by σ, we define $r^{\text{KMC}}(\sigma)$ as the steady state

Table 5.1: Elementary reactions in the A→B site-coupling model. The partial pressures of gas species A(g) and B(g) are $P_A = 1$ atm and $P_B = 0$ atm. The surface species are A, B, and empty sites (*). Subscripts indicate the site type as either terrace (t) or edge (e). Reactions 1-5 are adsorption, diffusion, surface reaction, diffusion, and desorption respectively. k_{fwd} and k_{rev} indicate forward and reverse rate constants, respectively.

Index	Reaction	k_{fwd}	$k_{\text{fwd}}/k_{\text{rev}}$
1	$\text{A(g)} \leftrightarrow \text{A}_t$	1.0 $\text{atm}^{-1}s^{-1}$	1.0 atm^{-1}
2	$\text{A}_t + *_t \leftrightarrow *_t + \text{A}_t$	1.0 s^{-1}	1.0
3	$\text{A}_t + *_e \leftrightarrow *_t + \text{B}_e$	1.0 s^{-1}	1.0
4	$\text{B}_e + *_e \leftrightarrow *_e + \text{B}_e$	1.0 s^{-1}	1.0
5	$\text{B}_e \leftrightarrow \text{B(g)} + *$	1.0 s^{-1}	1.0 atm

rate of production of a gas phase species P during KMC simulation, normalized by its stoichiometric coefficient in the net gas-phase reaction, as well as the size of the unit cell (n). The goal of structure optimization is to maximize $r^{\text{KMC}}(\sigma)$ with respect to σ. In theory, $r_{P,i}^{\text{KMC}}(\sigma)$ can be computed from the Chemical Master equation. Stochastic noise in the evaluation of $r^{\text{KMC}}(\sigma)$ from KMC simulation is managed by time averaging over sufficiently large time intervals. The Zacros[204, 148] graph theoretical software is used to simulate the reaction network. To ensure steady state conditions, we use the automated methods developed in Chapter 3.

For the purposes of extracting maximal data from each simulation, rates are computed for each Ni site. Given the steady state expected propensities of each elementary step, the rate of production of gas species P at site i is

$$r_{P,i}^{\text{KMC}}(\sigma) = \frac{1}{\nu_P} \sum_{j=1}^{n_{\text{rxns}}} \frac{\nu_{j,P}}{s_j} E\left[a_{i,j}\right]. \tag{5.1}$$

ν_P is the stoichiometric coefficient of species P in the net gas phase reaction. j indexes the elementary steps, of which there are n_{rxns}. $\nu_{j,P}$ is the stoichiometric coefficient of gas species P in elementary step j. s_j is the number of sites involved in elementary step j, which prevents double counting for reactions involving multiple sites. $E\left[a_{i,j}\right]$ denotes the steady state expectation of the propensity of elementary step j at Ni site i, as estimated from statistical sampling. If a Ni site i does not have a corresponding lattice site in the KMC simulation, then $r_{P,i}^{\text{KMC}}(\sigma) = 0$.

The total rate is computed as

$$r^{\mathrm{KMC}}(\sigma) = \frac{1}{n}\sum_{i=1}^{n} r_{P,i}^{\mathrm{KMC}}(\sigma). \tag{5.2}$$

Steady state mass balances render the total rate $r^{\mathrm{KMC}}(\sigma)$ independent of the choice of P. For our prototype reaction network, we use $P = B$ and $\nu_B = 1$ to track the production of B.

We exploit the symmetry of the (111) facet to maximize data usage from a single KMC calculation. For Ni site i, we define the translation operator $T_i(\sigma)$ as the permutation of σ that translates the catalyst so that Ni site i is mapped to $i = 1$. For $\theta \in \{0, 120, 240\}$, we define the rotational operator $R_\theta(\sigma)$ as the permutation of σ that rotates the catalyst by θ degrees counterclockwise about Ni site $i = 1$. For any combination of i, θ, P, and σ, symmetry dictates that

$$r_{P,1}^{KMC}\left(R_\theta\left(T_i\left(\sigma\right)\right)\right) = r_{P,i}^{KMC}(\sigma). \tag{5.3}$$

The right hand side of Equation 5.3 is known from simulation. Using KMC data for a single structure and all possible combinations of i and k gives $3n$ distinct mappings of σ to $r_{P,1}^{KMC}(\sigma)$. Further details on the Ni site indexing and symmetry operators are provided in Section D.1 of the supplement.

5.3.2 Surrogate model

A surrogate model ($r^{\mathrm{surr}}(\sigma)$) requiring low computational cost and sharing the same global maximum as $r^{\mathrm{KMC}}(\sigma)$ is needed for structure optimization. The surrogate model is built using machine learning tools implemented in the scikit-learn Python package[160]. The site rates ($r_{P,i}^{KMC}(\sigma)$), rather than the total rate ($r^{\mathrm{KMC}}(\sigma)$), are used as training data. Most Ni sites have zero rate, resulting in an unbalanced data set. Therefore, a combination of classification and regression is used. Figure 5.2 shows a diagram of the surrogate model.

First, a decision tree ($f^{\mathrm{DT}}(\sigma)$) classifies each site as either active or inactive. That is, $f^{\mathrm{DT}}(\sigma) = 1$ if $r_{P,1}^{\mathrm{surr}}(\sigma) > 0$ and $f^{\mathrm{DT}}(\sigma) = 0$ if $r_{P,1}^{\mathrm{surr}}(\sigma) = 0$. The input features

are the elements of the occupancy vector σ. The maximum depth of the tree is adjusted to maximize predictive accuracy, based on splitting the initial KMC database into a training and a test set.

A neural network is used to regress the rates of Ni sites with nonzero rate. Neural networks are tolerant to noise in the training data and tend to ignore irrelevant input features. This makes them suitable for our problem, given that the training data is evaluated using stochastic simulation and the occupancies of distant Ni occupancies tends not to affect the rate at a given Ni site. The structure of the neural network ($f^{\mathrm{NN}}(\sigma)$) consists of an input layer with one node for each occupancy in σ, 1 hidden layer, and an output node predicting $r^{\mathrm{surr}}_{P,1}$. The hidden layer nodes use a rectified linear unit (RELU) activation function while the output layer uses the identity function. The number of hidden nodes and the regularization parameter are chosen by minimizing the prediction error when regressed to the initial KMC database.

The site rates and the total rate are computed using the surrogate model as

$$r^{\mathrm{surr}}_{P,i}(\sigma) = f^{\mathrm{DT}}\left(T_i(\sigma)\right) * f^{\mathrm{NN}}\left(T_i(\sigma)\right) \tag{5.4}$$

$$r^{\mathrm{surr}}(\sigma) = \frac{1}{n}\sum_{i=1}^{n} r^{\mathrm{surr}}_{P,i}(\sigma). \tag{5.5}$$

Prediction of the total structure rate ($r^{\mathrm{surr}}(\sigma)$), being a spatial average of the site rates ($r^{\mathrm{surr}}_{P,i}(\sigma)$), is less noisy than for the individual site rates. However, this effect is counterbalanced by the larger noise in the KMC site rates ($r^{\mathrm{KMC}}_{P,i}(\sigma)$) used in the training data relative to the KMC structure rates ($r^{\mathrm{KMC}}(\sigma)$).

5.3.3 Simulated annealing-based optimization

For optimization of $r^{\mathrm{surr}}(\sigma)$, simulated annealing is employed. Each Metropolis step chooses an element of σ at random and attempts to flip it (1 to 0 or 0 to 1). The objective function is evaluated at each step using the surrogate model (Equation 5.5). The total number of steps ($s_{\max}$) and the cooling schedule are specified. An exponential

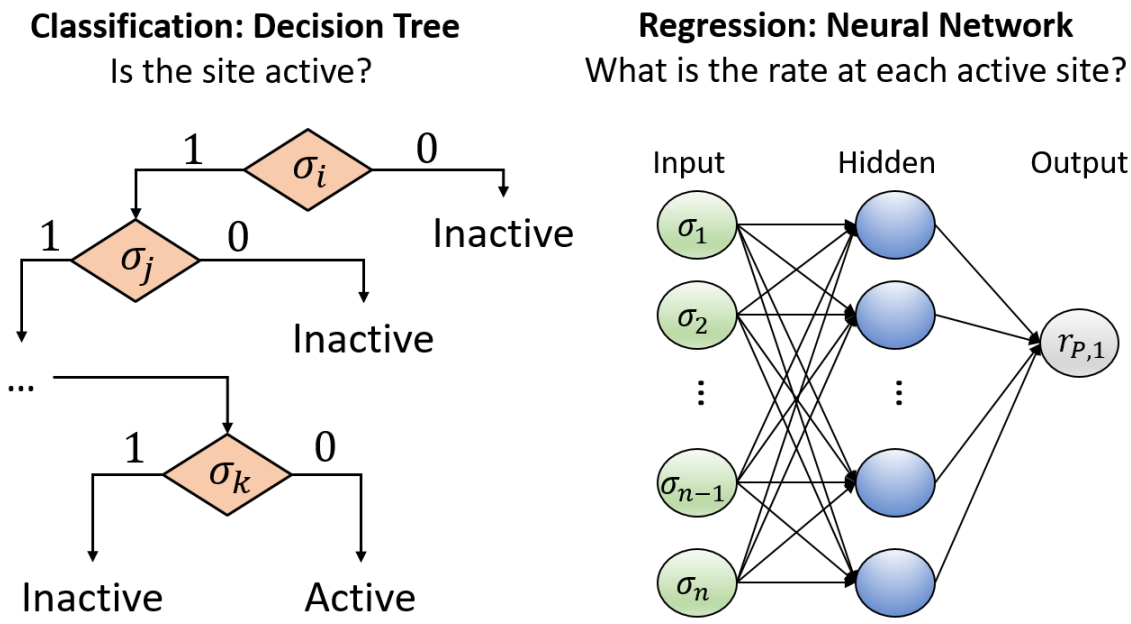


Figure 5.2: Diagram of the surrogate model.

cooling schedule is used, as it offers a suitable tradeoff between ease of implementation and performance[153]. The temperature $T(s)$ at step s is

$$T(s) = T_0 \exp\left(-\frac{s}{\tau}\right). \tag{5.6}$$

The parameter T_0 is the initial temperature and is set to the highest site rate observed in the training set, thus scaling the temperature by expected changes in the structure rate. The number of steps is determined based on the number of Ni sites, as n is the maximum number of Metropolis steps necessary to transition between any possible structure and the optimal one. We find that $s_{\max} = 100n = 14,400$ steps are sufficient in all cases. τ is the relaxation time scale chosen to be $\tau = \frac{1}{5}s_{\max}$. Judicious choices of the simulated annealing parameters ensure convergence to the global maximum of $r^{\text{surr}}(\sigma)$. For computational efficiency, a full list of all translations of the catalyst is maintained and updated to evaluate the summation in Equation 5.5.

5.3.4 Online machine learning (OML)

Initially, 50 structures are randomly generated by choosing a fixed Ni coverage and seeding Ni atoms at random locations, resulting in a variety of quantities and spatial arrangements of terrace and edge sites. KMC simulations are performed for the structures to create a KMC database of site and structure rates. Given that machine learning tools are highly interpolative, the surrogate model will converge to a false global optimum if the true optimum is dissimilar to the structures in the database[97].

To alleviate the issue of false convergence, optimization is performed iteratively using an OML approach. The predicted optimal structures are evaluated with KMC and added to the database, after which the surrogate model is retrained on the augmented database. Incremental learning is not employed for the neural network, given that catastrophic interference is likely to affect low activity sites of suboptimal structures. The refined surrogate model is used in subsequent simulated annealing optimization, using the previous optimum as the initial structure. The process is repeated until convergence, wherein the predicted optimal structure is evaluated correctly with the surrogate model and does not change with the addition of new data to the KMC database.

Parallelization is exploited to grow the database more rapidly. After evaluating the initial structures and adding them to the database, each of 16 processors perform the OML procedure independently. The processors communicate only by using data from the same database to which they all add. Use of multiple processors also makes the simulated annealing optimization more robust by increasing the probability of finding the global optimum. The OML process is shown in the flowchart in Figure 5.3 and summarized in the following pseudocode. Our software implementation is available at `https://github.com/VlachosGroup/Structure-Optimization`.

1. Generate initial training structures.
2. Evaluate training structures with KMC. Record site rates in a database.
3. Train the surrogate model on the structures in the database.

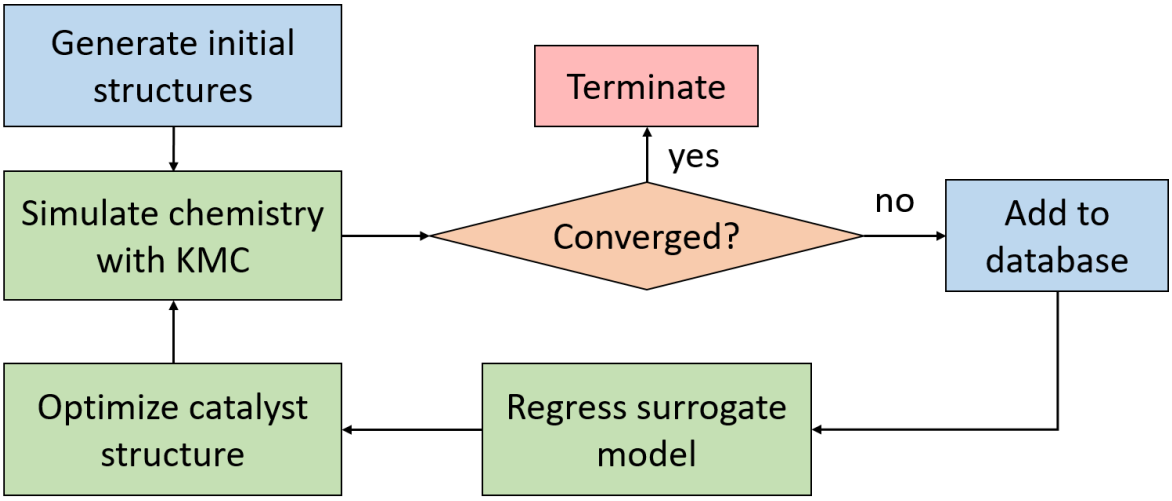


Figure 5.3: Flowchart of the online machine learning (OML) algorithm.

   a. Train decision tree to classify each site as having a zero or nonzero rate.
   b. Train neural network on all sites with nonzero rates.

4. Optimize catalyst structure using simulated annealing and the surrogate model ($r^{\text{surr}}(\sigma)$) as the objective function.

5. Evaluate predicted optimal structures with KMC, i.e., compute $r^{\text{KMC}}(\sigma)$.

6. If the optima have been evaluated correctly with the surrogate model, i.e., $r^{\text{surr}}(\sigma) \approx r^{\text{KMC}}(\sigma)$, and are the best historically, then terminate. Otherwise, add them to the KMC database and return to (3).

5.4 Results

5.4.1 Online machine learning

The maximum depth of the decision tree was chosen to be 20, on the basis of minimizing the prediction error when splitting the initial database into a 75% training and 25% test set. Figure 5.4 shows how the classification error is minimized for the test set when a maximum tree depth of 20 is used.

The parity plot for the neural network regression to individual site rates is shown in Figure 5.5. 20 hidden nodes were sufficient to achieve a reasonable fit to the training data. Variability in the site rates shows that they vary depending on the local microstructure.

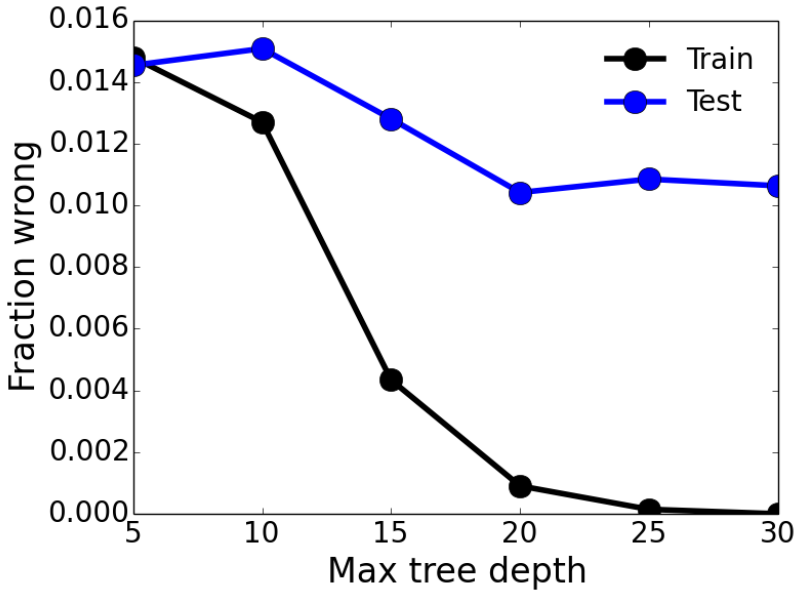


Figure 5.4: Classification error for the decision tree when trained to the initial database using different maximum depths.

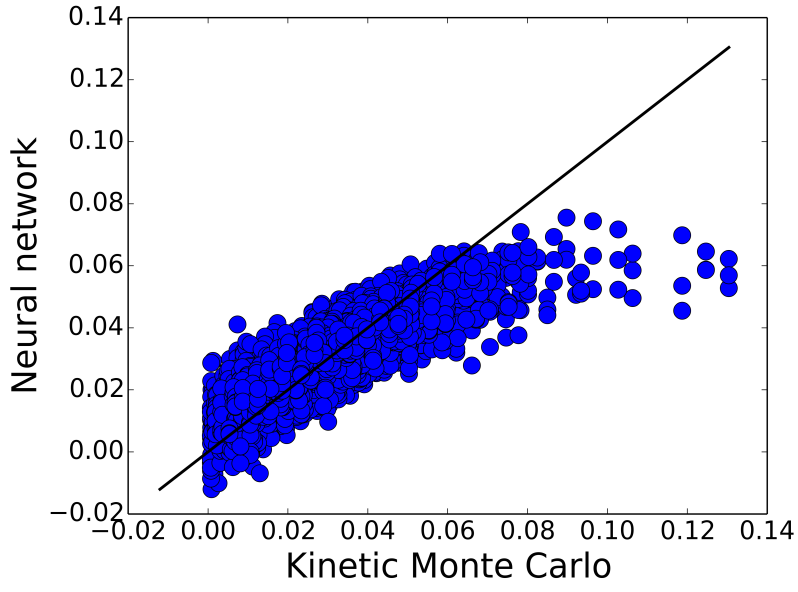


Figure 5.5: Parity plot of site rates ($r_{A,1}(\sigma)$), in units of molecules of B produced per second.

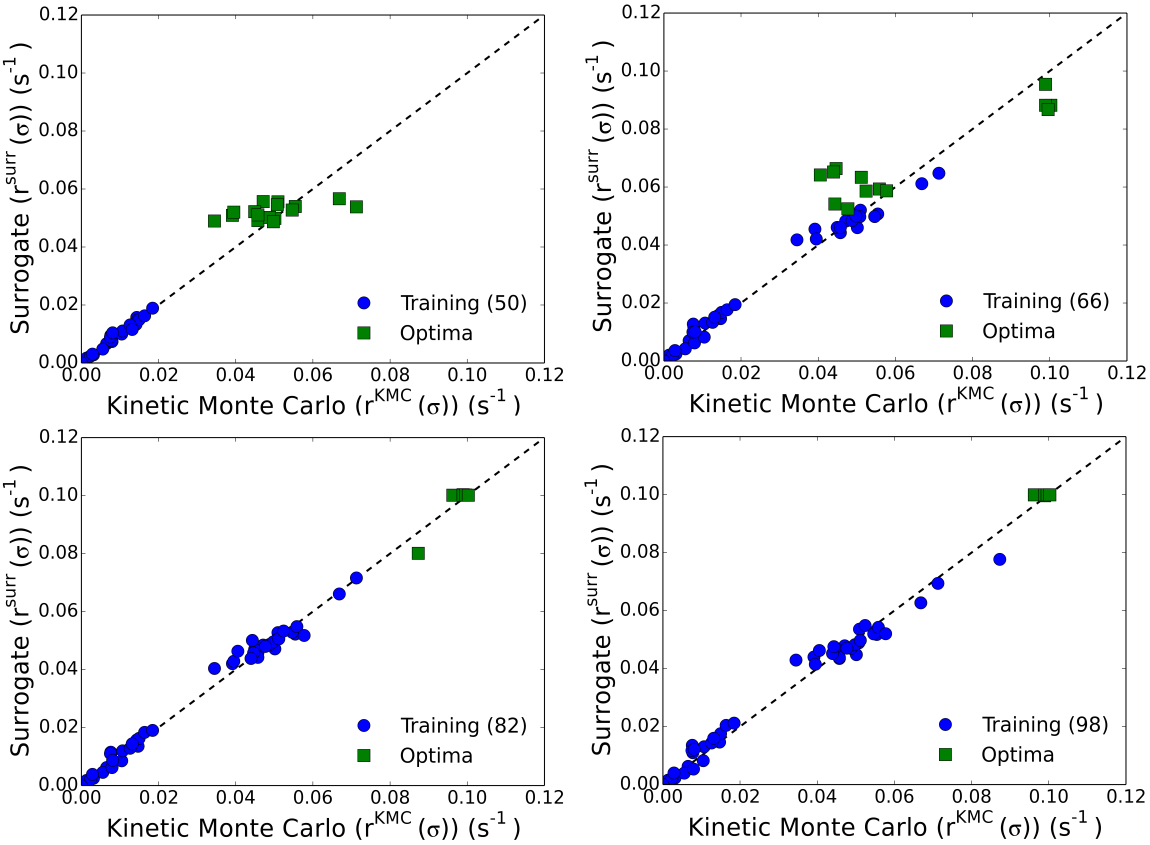


Figure 5.6: Parity plots of total structure rates ($r(\sigma)$) at the beginning (top left) and end (bottom right) of the online learning process. Green points are predicted by simulated annealing optimization using a surrogate model regressed to the data in the blue points.

The performance of the surrogate model in predicting catalyst activity is shown in Figure 5.6 for four iterations of the OML process, each of which add 16 structures to the database. In all iterations, simulated annealing optimization using the surrogate model finds the global maximum of $r^{\text{surr}}(\sigma)$, resulting in the green points always being higher than the blue points. In the first iteration, inaccuracies in the neural network cause the surrogate model to converge to false optima. The green points appear to the right of the blue points, but do not yet reach the global maximum of $r^{\text{KMC}}(\sigma)$. The surrogate model predictions improve after more data has been added to the database. By Iteration 4, in which 98 structures are used to regress the surrogate model, the surrogate model estimates the activity of the optimal structures accurately.

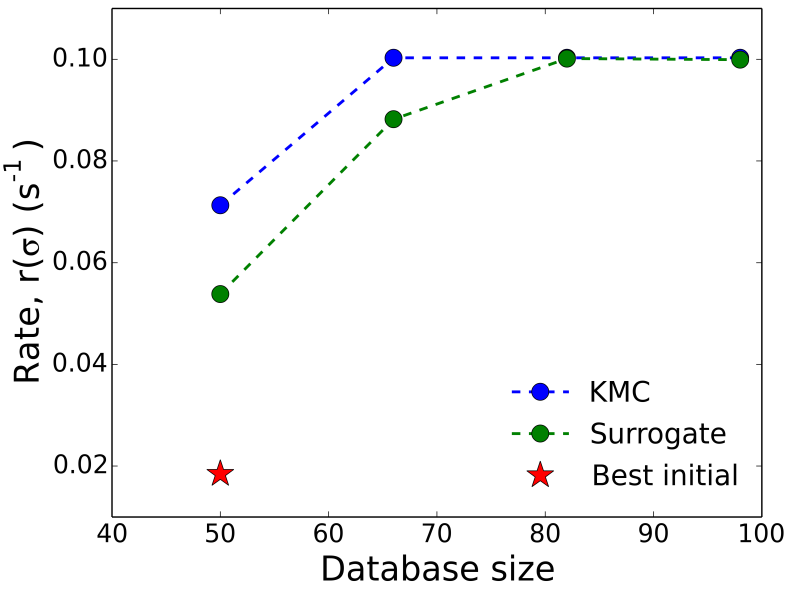


Figure 5.7: Surrogate model predictions ($r^{\text{surr}}(\sigma)$) and KMC evaluations ($r^{\text{KMC}}(\sigma)$) of the most active structure given by simulated annealing optimization, versus the number of structures in the KMC database. The red star indicates the most active structure in the initial KMC database as a point of comparison.

Figure 5.7 shows the convergence of the OML process with respect to the size of the KMC database. The surrogate model predictions and KMC evaluations of the most active structures predicted by simulated annealing are shown. As the size of the KMC database increases, the surrogate model achieves better agreement with KMC for the activity of the optimal structure. Eventually, the global optimum of $r^{\text{KMC}}(\sigma)$ is achieved. Relative to the most active structure in the initial KMC database, the global optimum has an activity higher by a factor of about 5. The OML process successfully discovers new and better structures starting only with information about random structures.

5.4.2 Physical insights from KMC simulation

The frequencies of terrace and edge sites in each structure are shown in Figure 5.8. Points are colored to indicate the activity of the corresponding structure, providing insights into how bifunctionality affects the catalyst structure-activity relationship.

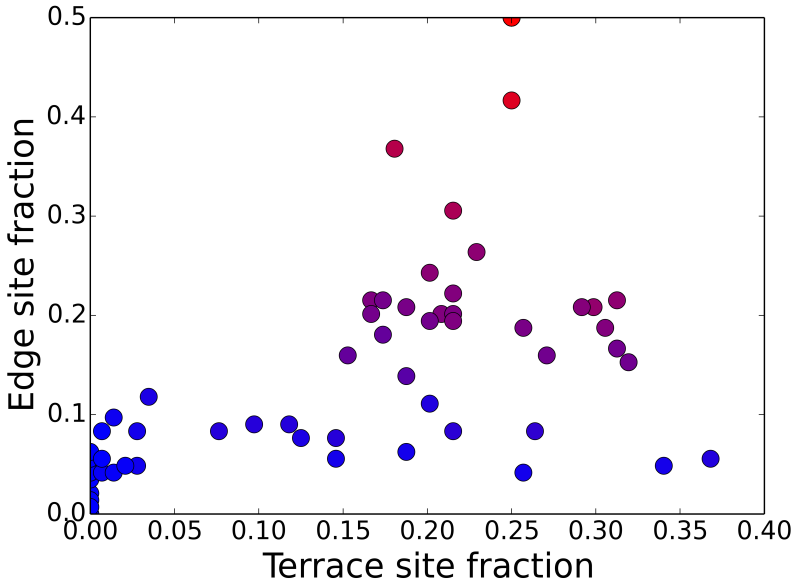


Figure 5.8: Fraction of Ni sites that are edge and terrace sites for all 114 structures encountered during optimization. Site fractions are normalized by the total number of Ni sites ($n = 144$). Because many Ni sites are not identified as either type of site, the fractions do not sum to 1. Markers are colored on a blue (least active) to red (most active) scale to qualitatively indicate the activity of each structure.

Although the activities seem to be controlled primarily by the amount of step sites, terrace sites are also necessary to achieve high activity.

Figure 5.9 shows the optimal structure, which exhibits several desirable properties. Not only does it have an abundance of both terrace and step sites, but it places them in close proximity so that after species A adsorbs onto a terrace site, it rapidly diffuses to an edge site where it desorbs as B.

5.4.3 CPU analysis

Figure 5.10 decomposes the central processing unit (CPU) requirements of the OML approach as well as a hypothetical brute force approach in which KMC is used at every step of optimization. The CPU cost of the brute force approach is purely due to KMC simulation, which must be performed at each of the 14,400 steps in the simulated annealing optimization. The OML approach requires 2000 times fewer KMC evaluations, but introduces the additional costs of regressing and evaluating the

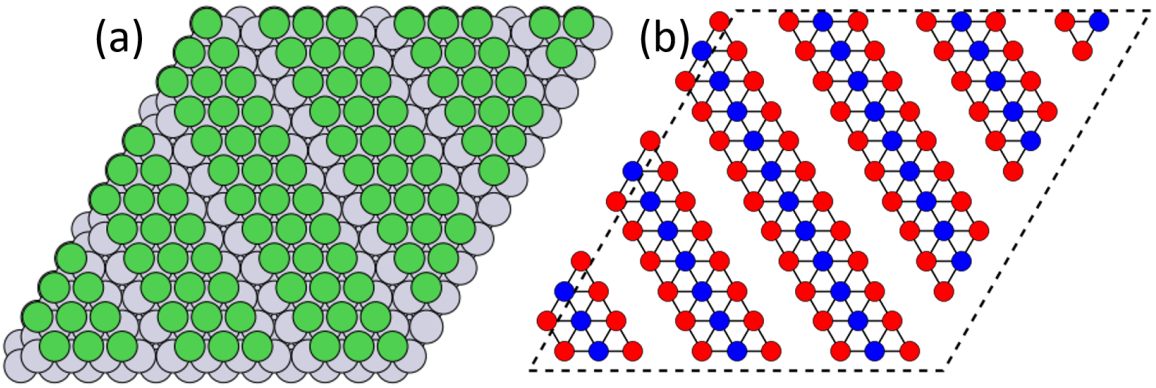


Figure 5.9: (a) Molecular picture of the catalyst structure with an optimal arrangement of a defected Ni adlayer (green) and a Pt substrate (grey). (b) The corresponding KMC lattice with terrace (blue) and edge (red) sites.

neural network. These extra tasks require far less CPU time than KMC evaluation does. The overall savings of the OML approach relative to brute force is a factor of 1600 for our example. Our CPU evaluations do not account for wall time reduction due to parallelization, as several strategies are available to control this. For example, a genetic algorithm would allow for parallel evaluation of the objective function for every individual in the population. The KMC simulations in this study require only 30 seconds to run, which is atypically fast for KMC. Typical simulations for applied problems usually require run times on the order of hours to days, in which case the CPU savings of the OML approach would be even more significant.

5.5 Conclusions

An approach to computationally optimizing catalyst structure for chemistries with coupled active sites was developed. A surrogate model consisting of a decision tree and a neural network was regressed to data from KMC simulation of various structures. The surrogate model identified active sites and predicted their activity with sufficient accuracy to be used in simulated annealing optimization. An iterative approach was used in which predicted optimal structures were evaluated with KMC and added to a database to refine the surrogate model.

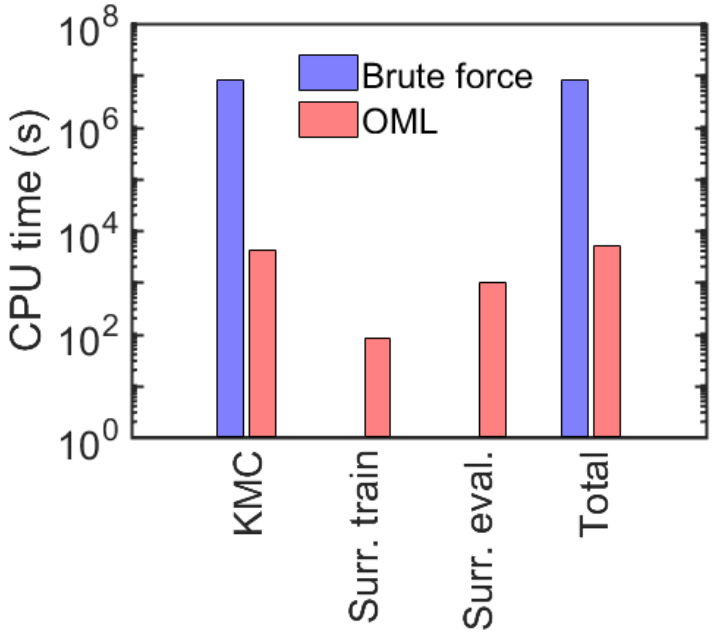


Figure 5.10: CPU comparison of the brute force approach (projected cost of using KMC directly in optimization) versus our online machine learning (OML) approach. Relative to brute force, OML requires significantly fewer KMC simulations, but introduces the additional costs of training (train NN) and evaluating the neural network (NN eval.).

We applied our technique to a prototype chemistry exhibiting bifunctional behavior between terrace and step sites of defected Ni adlayers on a Pt substrate. The accuracy of the surrogate model improved as more structures were added to the KMC database. Simulated annealing optimization predicted an optimal structure with an abundance of terrace and step sites in close proximity to each other. Subsequent KMC evaluation confirmed its activity to be 5 times greater than the best structure in the initial set of randomly generated training structures. Use of a surrogate model in optimization reduced the number of KMC evaluations necessary by a factor of 2000 relative to a hypothetical brute force scheme.

Our new methodology shows promise for designing more active catalysts for a variety of structure sensitive chemistries. Core-shell catalysts as well as metals on oxide supports are commonly used catalysts and expose a variety of different sites which affect the chemistry differently. Computational methods for estimating binding properties at different sites continue to be developed. Once optimal structures are predicted, experimental efforts can be made to synthesize them, thus providing crucial benchmarking for structure dependent kinetic models. Our optimization approach mapping catalyst structure to activity can be used in tandem with experimental characterization of catalyst structure using techniques such as X-ray absorption spectroscopy (XAS)[189] and extended x-ray absorption fine-structure spectroscopy (EXAFS)[177].

Chapter 6

SUMMARY AND OUTLOOK

In this chapter, we summarize the accomplishments of this thesis and outline directions for future research.

6.1 Dissertation Conclusions

This thesis developed several uncertainty quantification methods for kinetic Monte Carlo (KMC) simulation of catalytic reactions. Chapters 2 and 3 developed sensitivity analysis methods for quantifying uncertainty due to errors in model parameters. Chapters 4 and 5 developed catalyst structure optimization methods to assess the impact of catalyst structure and the identity of the active site.

Chapter 2 extended the likelihood ratio sensitivity analysis method, an efficient alternative to finite differencing, to multiscale KMC simulations. For a prototype chemical reaction network with reactions occurring on different time scales, the variance of likelihood ratio sensitivity estimates for parameters affecting the fast events was shown to be unmanageable. To circumvent this limitation of the likelihood ratio method, the time scales of the reaction network were explicitly separated into fast and slow modes. The fast time scale was equilibrated, whereas the slow scale was simulated with KMC using propensities determined by averaging on the fast time scale. Sensitivity contributions were computed separately for each time scale and combined to achieve an overall sensitivity with drastically lower variance than the single time scale version.

Chapter 3 developed statistical techniques to accelerate lattice KMC simulation and enable likelihood ratio sensitivity analysis for realistic chemistry, for which unknown time scales challenge statistical sampling. A prototype A$\rightarrow$B model and

the water-gas shift reaction on Pt(111) were used as test cases. A robust criterion for determining non-equilibrium steady state conditions was developed based on batch means. Parallel processing and rate constant rescaling significantly accelerated KMC simulation without affecting the slow dynamics of the system, with the latter being especially significant. Likelihood ratio sensitivity analysis using time averaged properties correctly identified the rate determining step in the water-gas shift model as dissociation of adsorbed water, as benchmarked against finite differencing. Likelihood ratio sensitivity analysis was shown to be faster than finite differencing for the water-gas shift system, but not for the A→B model, due to CPU dependence on the number of reactions of a network.

Chapter 4 developed an optimization approach to maximizing catalyst activity by identifying the active site and the density and spatial arrangement of such sites while simultaneously minimizing surface energy. This methodology was applied to the oxygen reduction reaction (ORR) on defected Pt(111), Pt(100), Au(111), and Au(100) surfaces. The activity of each site was computed according to its generalized coordination number, which describes the site's local environment. We showed that not all sites can have optimal properties due to geometric and stability constraints on the low index planes. Our results rationalize why in contrast to optimistic theoretical predictions, only modest activity increases are seen in experimental studies. The optimal arrangements of active sites predicted by our new method are promising for further increasing the activity of catalyst surfaces with purposefully engineered defects.

Chapter 5 enabled catalyst structure optimization to account for active-site coupling originating from complex phenomena such as bifunctionality, adsorbate interactions, and diffusion. We demonstrated our method for a prototype chemistry exhibiting bifunctionality. Because the computational demands of KMC make it unsuitable for direct use in optimization, we constructed a surrogate model using neural networks regressed to KMC data. The neural networks captured the complex structural dependencies of KMC at comparably negligible computational cost. Because inaccuracies in the surrogate model caused the optimization to converge to false optima, an online

machine learning (OML) approach was used. Structures predicted by the optimization were evaluated using the KMC simulation and added to a database of structures. The surrogate model was retrained and the optimization was repeated in an iterative fashion. In doing so, the true global optimum was eventually identified and predicted accurately by the surrogate model. The online machine learning approach reduced the computational cost of optimization by a factor of 3 orders of magnitude relative to a brute force approach. The optimal structure contained an abundance of the two types of sites involved in the chemistry, located in close proximity to facilitate diffusion. Our findings offer new insights for designing bifunctional catalysts.

6.2 Future Directions

To further facilitate progress towards the eventual goal of robust first principles based catalyst design, we suggest directions for future research in each of the two major areas of this thesis.

6.2.1 Sensitivity analysis for kinetic Monte Carlo simulation

As shown in Chapter 3, the appropriate choice of local sensitivity analysis method depends on the system being simulated. Therefore, various methods should still be developed along with guidelines for choosing the most suitable technique for a given model. Recent advances have become increasingly focused on lattice KMC simulation[7, 121], making them usable for catalysis models. In some cases, a hybrid approach[87] will be the best choice. Given the advancements in this thesis and elsewhere, local sensitivity analysis could become routine for KMC studies of catalytic systems.

Global sensitivity analysis, described in Section 1.3, must also be investigated for KMC simulation. Previous work in this area has focused on biological models[103, 181]. Stochastic coupling[6, 199, 7], a method currently used to reduce the variance of finite differencing by introducing correlations between trajectories with different parameters, may play a similar role in global sensitivity analysis. Correlated sensitivity analysis

has been shown to affect interpretations of mean field models[210] and should also be performed for KMC models.

An important input for KMC is the cluster expansion Hamiltonian, which is fit to density functional theory (DFT) data for many different adsorbate configurations. Although error analysis has been performed for cluster expansion predictions of configuration energies[239, 141, 190], it is unknown how this error propagates through KMC simulation. An understanding of which DFT configurations affect the calculation of KMC observables would allow cluster expansion training to be restricted to DFT data that affects the KMC observables. Sensitivity analysis on cluster expansion coefficients is difficult because these parameters affect both fast and slow reactions.

6.2.2 Catalyst structure optimization

Catalyst structure optimization relies heavily on the relationships used to estimate active site properties as a function of their local geometry. These relationships should be pursued further to improve their accuracy and extend their applicability. For example, structure dependent linear scaling relationships[28] would enable the prediction of binding energies on different metals and sites simultaneously. Similarly, lateral interactions that scale across different structures would make structure dependent models more robust. In practice, this may take the form of a Hamiltonian cluster expansion with coordination dependent terms. The development of these relationships will involve a combination of fundamentals and DFT data. In the fundamental approach, analysis of the effect of catalyst structure on binding energies through the metal d-band[77, 142, 88, 29] would help derive physically motivated relationships. With a DFT approach, machine learning[221, 19] could be employed.

Once optimal structures are predicted, experimental efforts should be made to synthesize them. Experimental validation would provide crucial benchmarking for structure dependent kinetic models. Given the results of Chapter 4 for the oxygen reduction reaction, synthesis efforts should aim to maximize the quantity of cavities on the Pt(111) surface.

Several extensions of our machine learning approach to structure optimization can be made. It can be generalized to three dimensional nanoparticles to more closely resemble supported nanoparticles encountered in the laboratory. Running KMC simulations at different temperatures and partial pressures and incorporating these variables into to the neural network regression would enable optimization of catalyst structure and reaction conditions simultaneously. Our machine learning approach to mapping catalyst structure to activity would also enable simulations with a dynamic catalyst surface for which the KMC lattice changes with time.

6.3 Closing thoughts

As is the case for hurricane modeling and election forecasting, predictive models for catalyst design require uncertainty quantification to assess the level of confidence in model predictions. As theory and hardware improve, this confidence will gradually increase. Quantum computing, for example, shows great promise for performing the types of parallel calculations ubiquitous in scientific research. The construction of robust models will require a highly collaborative, multidisciplinary effort in addressing phenomena at different length and time scales. Data sharing and open-source software will facilitate such collaboration. To this end, links to software implementations of the methods developed in this thesis have been provided in the relevant chapters.

BIBLIOGRAPHY

[1] Technology Roadmap: Energy and GHG Reductions in the Chemical Industry via Catalytic Processes. Technical report, International Energy Agency, 2013.

[2] Annual Energy Outlook 2017 with projections to 2050. Technical report, U.S. Energy Information Administration, 2017.

[3] Frank Abild-Pedersen, Jeff Greeley, Felix Studt, Jan Rossmeisl, T R Munter, Poul Georg Moses, Egill Skulason, Thomas Bligaard, and Jens Kehlet Nørskov. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Physical Review Letters*, 99(1):16105, 2007.

[4] Hristiyan A Aleksandrov, Lyudmila V Moskaleva, Zhi-Jian Zhao, Duygu Basaran, Zhao-Xu Chen, Donghai Mei, and Notker Rösch. Ethylene conversion to ethylidyne on Pd(1 1 1) and Pt(1 1 1): A first-principles-based kinetic Monte Carlo study. *Journal of Catalysis*, 285(1):187–195, 2012.

[5] Mie Andersen, Craig P. Plaisance, and Karsten Reuter. Assessment of mean-field microkinetic models for CO methanation on stepped metal surfaces using accelerated kinetic Monte Carlo. *The Journal of Chemical Physics*, 147(15):152705–152718, oct 2017.

[6] David F Anderson. An efficient finite difference method for parameter sensitivities of continuous time Markov chains. *SIAM Journal on Numerical Analysis*, 50(5):2237–2258, 2012.

[7] Georgios Arampatzis and Markos A Katsoulakis. Goal-oriented sensitivity analysis for lattice kinetic Monte Carlo simulations. *The Journal of chemical physics*, 140(12):124108, 2014.

[8] Georgios Arampatzis, Markos A. Katsoulakis, and Luc Rey-Bellet. Efficient estimators for likelihood ratio sensitivity indices of complex stochastic dynamics. *Journal of Chemical Physics*, 144(10):104107, 2016.

[9] Giorgos Arampatzis, Markos A Katsoulakis, and Petr Plechac. Parallelization, processor communication and error analysis in lattice kinetic Monte Carlo. *SIAM Journal on Numerical Analysis*, 52(3):1156–1182, 2014.

[10] Sara Aranifard, Salai Cheettu Ammal, and Andreas Heyden. On the importance of metal-oxide interface sites for the water-gas shift reaction over Pt/CeO2 catalysts. *Journal of Catalysis*, 309:314–324, 2014.

[11] Adam Arkin, John Ross, and Harley H McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics*, 149(4):1633–1648, 1998.

[12] John N. Armor. A history of industrial catalysis. *Catalysis Today*, 163(1):3–9, apr 2011.

[13] Ashish Arora and Alfonso Gambardella. Implications for Energy Innovation from the Chemical Industry. In *Accelerating Energy Innovation: Insights from Multiple Sectors*, pages 87–111. University of Chicago Press, may 2011.

[14] Bernd A Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*. World Scientific, Singapore, oct 2004.

[15] D Wayne Blaylock, Yi-An Zhu, and William H Green. Computational investigation of the thermochemistry and kinetics of steam methane reforming over a multi-faceted nickel catalyst. *Topics in Catalysis*, 54(13-15):828–844, 2011.

[16] T Bligaard, Jens Kehlet Norskov, Soren Dahl, J Matthiesen, Claus H Christensen, and J Sehested. The BrønstedEvansPolanyi relation and the volcano curve in heterogeneous catalysis. *Journal of Catalysis*, 224(1):206–217, may 2004.

[17] Thomas Bligaard, R. Morris Bullock, Charles T. Campbell, Jingguang G. Chen, Bruce C. Gates, Raymond J. Gorte, Christopher W. Jones, William D. Jones, John R. Kitchin, and Susannah L. Scott. Toward Benchmarking in Catalysis Science: Best Practices, Challenges, and Opportunities. *ACS Catalysis*, 6(4):2590–2602, apr 2016.

[18] Peter E Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953, 1994.

[19] Jacob R. Boes and John R. Kitchin. Neural network predictions of oxygen interactions on a dynamic Pd surface. *Molecular Simulation*, 43(5-6):346–354, apr 2017.

[20] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *Journal of Computational Physics*, 17(1):10–18, 1975.

[21] Mary Ellen Bowden and John Kenly Smith. *American Chemical Enterprize*. Chemical Heritage Foundation, Philadelphia, PA, 1994.

[22] Kaitlin M Bratlie, Hyunjoo Lee, Kyriakos Komvopoulos, Peidong Yang, and Gabor A Somorjai. Platinum nanoparticle shape effects on benzene hydrogenation selectivity. *Nano letters*, 7(10):3097–3101, 2007.

[23] Felix Brockherde, Leslie Vogt, Li Li, Mark E. Tuckerman, Kieron Burke, and Klaus-Robert Müller. Bypassing the Kohn-Sham equations with machine learning. *Nature Communications*, 8(1):872, dec 2017.

[24] P C Bruijnincx and B M Weckhuysen. Shale gas revolution: an opportunity for the production of biobased chemicals? *Angewandte Chemie. International Ed. In English*, 52(46):11980–11987, 2013.

[25] Stephen Brunauer, P. H. Emmett, and Edward Teller. Adsorption of Gases in Multimolecular Layers. *Journal of the American Chemical Society*, 60(2):309–319, feb 1938.

[26] Qiuxia Cai, Xinde Wang, and Jian-guo Wang. Distinctions between supported Au and Pt catalysts for CO oxidation: insights from DFT study. *The Journal of Physical Chemistry C*, 117(41):21331–21336, 2013.

[27] Federico Calle-Vallejo and Marc T M Koper. First-principles computational electrochemistry: Achievements and challenges. *Electrochimica Acta*, 84:3–11, 2012.

[28] Federico Calle-Vallejo, David Loffreda, Marc T. M. Koper, and Philippe Sautet. Introducing structural sensitivity into adsorptionenergy scaling relations by means of coordination numbers. *Nature Chemistry*, 7(5):403–410, 2015.

[29] Federico Calle-Vallejo, José I. Martínez, Juan M. García-Lastra, Philippe Sautet, and David Loffreda. Fast prediction of adsorption properties for platinum nanocatalysts with generalized coordination numbers. *Angewandte Chemie - International Edition*, 53(32):8316–8319, 2014.

[30] Federico Calle-Vallejo, Marcus D Pohl, David Reinisch, David Loffreda, Philippe Sautet, and Aliaksandr S Bandarenka. Why conclusions from platinum model surfaces do not necessarily lead to enhanced nanoparticle catalysts for the oxygen reduction reaction. *Chemical Science*, 2017.

[31] Federico Calle-Vallejo, Jakub Tymoczko, Viktor Colic, Quang Huy Vu, Marcus D. Pohl, Karina Morgenstern, David Loffreda, Philippe Sautet, Wolfgang Schuhmann, and Aliaksandr S. Bandarenka. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science*, 350(6257):185–189, 2015.

[32] Charles T. Campbell. Finding the Rate-Determining Step in a Mechanism: Comparing DeDonder Relations with The "Degree of Rate Control". *Journal of Catalysis*, 204(2):520–524, 2001.

[33] Y Cao, D T Gillespie, and L R Petzold. Avoiding negative populations in explicit Poisson tau-leaping. *Journal of Chemical Physics*, 123(5):54104, 2005.

[34] Y Cao, D T Gillespie, and L R Petzold. Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*, 124:44109, 2006.

[35] Yang Cao, Dan Gillespie, and Linda Petzold. Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems. *Journal of Computational Physics*, 206(2):395–411, 2005.

[36] Yang Cao, Daniel T. Gillespie, and Linda R. Petzold. The slow-scale stochastic simulation algorithm. *Journal of Chemical Physics*, 122(1):14116, 2005.

[37] Giuliano Carchini, Neyvis Almora-Barrios, Guillem Revilla-López, Luca Bellarosa, Rodrigo García-Muelas, Max García-Melchor, Sergey Pogodin, Piotr Błoński, and Núria López. How Theoretical Simulations Can Address the Structure and Activity of Nanoparticles. *Topics in Catalysis*, 56(13-14):1262–1272, 2013.

[38] Matteo Cargnello, Vicky V. T. Doan-Nguyen, Thomas R. Gordon, Rosa E. Diaz, Eric A. Stach, Raymond J. Gorte, Paolo Fornasiero, and Christopher B. Murray. Control of Metal Nanocrystal Size Reveals Metal-Support Interface Role for Ceria Catalysts. *Science*, 341(6147):771–773, 2013.

[39] Anisha Chakrabarti, Michael E Ford, Daniel Gregory, Rongrong Hu, Christopher J Keturakis, Soe Lwin, Yadan Tang, Zhou Yang, Minghui Zhu, Miguel A Bañares, and Israel E. Wachs. A decade+ of operando spectroscopy studies. *Catalysis Today*, 283:27–53, apr 2017.

[40] A Chatterjee, D G Vlachos, and M A Katsoulakis. Binomial distribution based tau-leap accelerated stochastic simulation. *Journal of Chemical Physics*, 122(2):24112, 2005.

[41] Abhijit Chatterjee and Dionisios G. Vlachos. An overview of spatial microscopic and accelerated kinetic Monte Carlo methods. *Journal of Computer-Aided Materials Design*, 14(2):253–308, 2007.

[42] Abhijit Chatterjee and Arthur F. Voter. Accurate acceleration of kinetic Monte Carlo simulations through the modification of rate constants. *Journal of Chemical Physics*, 132(19):194101, 2010.

[43] I. Chorkendorff and J. W. Niemantsverdriet. *Concepts of modern catalysis and kinetics*. John Wiley & Sons., 2003.

[44] Rune Christensen, Heine A. Hansen, Colin F. Dickens, Jens K. Nørskov, and Tejs Vegge. Functional Independent Scaling Relation for ORR/OER Catalysts. *Journal of Physical Chemistry C*, 120(43):24910–24916, 2016.

[45] Phillip Christopher and Suljo Linic. Engineering selectivity in heterogeneous catalysis: Ag nanowires as selective ethylene epoxidation catalysts. *Journal of the American Chemical Society*, 130(34):11264–11265, 2008.

[46] Piotr Czyżżak and Adrezej Jaszkiewicz. Pareto simulated annealinga metaheuristic technique for multipleobjective combinatorial optimization. *Journal of MultiCriteria Decision Analysis*, 7(1):34–47, 1998.

[47] Thomas Danielson, Jonathan E Sutton, Céline Hin, and Aditya Savara. SQERTSS: Dynamic rank based throttling of transition probabilities in kinetic Monte Carlo simulations. *Computer Physics Communications*, 219:149–163, 2017.

[48] William M Deen. *Analysis of Transport Phenomena (Topics in Chemical Engineering)*, volume 3. Oxford University Press, New York, 1998.

[49] John R. Di Iorio, Claire T. Nimlos, and Rajamani Gounder. Introducing Catalytic Diversity into Single-Site Chabazite Zeolites of Fixed Composition via Synthetic Control of Active Site Proximity. *ACS Catalysis*, 7(10):6663–6674, oct 2017.

[50] Laetitia Dubau, Jaysen Nelayah, Simona Moldovan, Ovidiu Ersen, Pierre Bordet, Jakub Drnec, Tristan Asset, Raphaël Chattot, and Frédéric Maillard. Defects do Catalysis: CO Monolayer Oxidation and Oxygen Reduction Reaction on Hollow PtNi/C Nanoparticles. *ACS Catalysis*, 6(7):4673–4684, 2016.

[51] Eric C Dybeck, Craig P Plaisance, and Matthew Neurock. Generalized Temporal Acceleration Scheme for Kinetic Monte Carlo Simulations of Surface Catalytic Processes by Scaling the Rates of Fast Reactions. *Journal of Chemical Theory and Computation*, 13:1525–1538, 2017.

[52] Weinan E, Di Liu, and Eric Vanden-Eijnden. Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales. *Journal of Computational Physics*, 221(1):158–180, 2007.

[53] A Eichler, J Hafner, J Furthmüller, G Kresse, J Furthmuller, and G Kresse. Structural and electronic properties of rhodium surfaces: an ab initio approach. *Surface Science*, 346(1):300–321, 1996.

[54] Jan Willem Erisman, Mark A. Sutton, James Galloway, Zbigniew Klimont, and Wilfried Winiwarter. How a century of ammonia synthesis changed the world. *Nature Geoscience*, 1(10):636–639, oct 2008.

[55] Hanne Falsig, Britt Hvolbæk, Iben S Kristensen, Tao Jiang, Thomas Bligaard, Claus H Christensen, and Jens K Nørskov. Trends in the catalytic CO oxidation activity of nanoparticles. *Angewandte Chemie*, 120(26):4913–4917, 2008.

[56] Peter Ferrin and Manos Mavrikakis. Structure sensitivity of methanol electrooxidation on transition metals. *Journal of the American Chemical Society*, 131(40):14381–14389, 2009.

[57] Kristen A. Fichthorn and Yangzheng Lin. A local superbasin kinetic Monte Carlo method. *Journal of Chemical Physics*, 138(16):164104, 2013.

[58] Hubert A Gasteiger, Shyam S Kocha, Bhaskar Sompalli, and Frederick T Wagner. Activity benchmarks and requirements for Pt, Pt-alloy, and non-Pt oxygen reduction catalysts for PEMFCs. *Applied Catalysis B: Environmental*, 56(1):9–35, 2005.

[59] Aytekin Gel, Tingwen Li, Balaji Gopalan, Mehrdad Shahnam, and Madhava Syamlal. Validation and Uncertainty Quantification of a Multiphase Computational Fluid Dynamics Model. *Industrial & Engineering Chemistry Research*, 52:11424–11435, 2013.

[60] G. Germani and Y. Schuurman. Water-gas shift reaction kinetics over μ-structured Pt/CeO2/Al2O3 catalysts. *AIChE Journal*, 52(5):1806–1813, may 2006.

[61] D T Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81:2340–2361, 1977.

[62] D T Gillespie. The chemical Langevin equation. *Journal of Chemical Physics*, 113(1):297–306, 2000.

[63] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.

[64] Daniel T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1-3):404–425, 1992.

[65] Daniel T Gillespie. *Markov Processes: An Introduction for Physical Scientists*. Academic Press, Inc., Sand Diego, CA, 1992.

[66] Peter W Glynn. Optimization of Stochastic Systems via Simulation. In E.A. MacNair, P. Heidelberger, and K.J. Musselman, editors, *Proceedings of the 1989 Winter Simulation Conference*, pages 90–105, New York, NY, 1989. Association for Computing Machinery.

[67] Peter W. Glynn. Likelihood ratio gradient estimation for stochastic systems. *Commun. ACM*, 33(10):75–84, 1990.

[68] John Goutsias. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *Journal of Chemical Physics*, 122(18):184102, 2005.

[69] Geun Ho Gu and Dionisios G. Vlachos. Group Additivity for Thermochemical Property Estimation of Lignin Monomers on Pt(111). *The Journal of Physical Chemistry C*, 120(34):19234–19241, sep 2016.

[70] Samuel Guerin, Brian E Hayden, Derek Pletcher, Michael E Rendall, and Jens-Peter Suchsland. A combinatorial approach to the study of particle size effects on supported electrocatalysts: Oxygen reduction on gold. *Journal of combinatorial chemistry*, 8(5):679–686, 2006.

[71] Rudiyanto Gunawan, Yang Cao, Linda Petzold, and Francis J Doyle III. Sensitivity Analysis of Discrete Stochastic Systems. *Biophysical Journal*, 88(4):2530–2540, 2005.

[72] W Guo and D G Vlachos. On factors controlling activity of submonolayer bimetallic catalysts: Nitrogen desorption. *Journal of Chemical Physics*, 140(1):014703–014709, 2014.

[73] Wei Guo, Michail Stamatakis, and Dionisios G Vlachos. Design Principles of Heteroepitaxial Bimetallic Catalysts. *ACS Catalysis*, 3(10):2248–2255, 2013.

[74] Wei Guo and Dionisios G. Vlachos. Effect of local metal microstructure on adsorption on bimetallic surfaces: Atomic nitrogen on NiPt(111). *Journal of Chemical Physics*, 138(17), 2013.

[75] Wei Guo and Dionisios G Vlachos. Patched bimetallic surfaces are active catalysts for ammonia decomposition. *Nature communications*, 6:8619–8625, 2015.

[76] Ankit Gupta and Mustafa Khammash. Sensitivity analysis for stochastic chemical reaction networks with multiple time-scales. *Electronic Journal of Probability*, 19(59):1–53, 2014.

[77] B. Hammer and J.K. Norskov. Theoretical Surface Science and Catalysis Calculations and Concepts. *Advances in Catalysis*, 45:71–129, 2000.

[78] Christopher L Hanselman and Chrysanthos E Gounaris. A mathematical optimization framework for the design of nanopatterned surfaces. *AIChE Journal*, 62(9):3250–3263, 2016.

[79] Danielle A. Hansgen, Dionisios G. Vlachos, and Jingguang G. Chen. Using first principles to predict bimetallic catalysts for the ammonia decomposition reaction. *Nature Chemistry*, 2(6):484–489, 2010.

[80] CK Harris, D Roekaerts, FJJ Rosendal, FGJ Buitendijk, Ph Daskopoulos, AJN Vreenegoor, and H Wang. Computational fluid dynamics for chemical reactor engineering. *Chemical Engineering Science*, 51(10):1569–1594, may 1996.

[81] Eric L. Haseltine and James B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *Journal of Chemical Physics*, 117(15):6959–6969, 2002.

[82] Araz Hashemi, Marcel Núñez, Petr Plecháč, and Dionisios G. Vlachos. Stochastic averaging and sensitivity analysis for two scale reaction networks. *Journal of Chemical Physics*, 144(7):074104, 2016.

[83] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22):9901–9904, nov 2000.

[84] HenrikTopsøe. Developments in operando studies and in situ characterization of heterogeneous catalysts. *Journal of Catalysis*, 216(1-2):155–164, may 2003.

[85] Jeffrey A Herron, Jessica Scaranto, Peter Ferrin, Sha Li, and Manos Mavrikakis. Trends in formic acid decomposition on model transition metal surfaces: A density functional theory study. *ACS Catalysis*, 4(12):4434–4445, 2014.

[86] Ask Hjorth Larsen, Jens JØrgen Mortensen, Jakob Blomqvist, Ivano E. Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N. Groves, BjØrk Hammer, Cory Hargus, Eric D. Hermes, Paul C. Jennings, Peter Bjerre Jensen, James Kermode, John R. Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob SchiØtz, Ole Schütt, Mikkel Strange, Kristian S. Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W. Jacobsen. The atomic simulation environment - A Python library for working with atoms. *Journal of Physics Condensed Matter*, 29(27):273002–273031, jun 2017.

[87] Max J. Hoffmann, Felix Engelmann, and Sebastian Matera. A practical approach to the sensitivity analysis for kinetic Monte Carlo simulation of heterogeneous catalysis. *The Journal of Chemical Physics*, 146(4):044118, 2017.

[88] Max J. Hoffmann, Andrew J. Medford, and Thomas Bligaard. Framework for Scalable Adsorbate-Adsorbate Interaction Models. *Journal of Physical Chemistry C*, 120(24):13087–13094, 2016.

[89] K Honkala, Anders Hellman, I N Remediakis, Ashildur Logadottir, A Carlsson, Søren Dahl, Claus H Christensen, and Jens Kehlet Nørskov. Ammonia synthesis from first-principles calculations. *Science*, 307(5709):555–558, 2005.

[90] George W Huber, Sara Iborra, and Avelino Corma. Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering. *Chemical Reviews*, 106(9):4044–4098, 2006.

[91] M. Z. Jacobson, W. G. Colella, and D. M. Golden. Cleaning the Air and Improving Health with Hydrogen Fuel-Cell Vehicles. *Science*, 308(5730), 2005.

[92] Mitch Jacoby. Fuel-cell cars finally drive off the lot. *C&EN Global Enterprise*, 95(38):28–32, sep 2017.

[93] Tobias Jahnke, Wilhelm Huisinga, T Jahnke, and W Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *J. Math. Biol*, 54(1):1–26, 2007.

[94] A. P. J. (Antonius Petrus Johannes) Jansen. *An introduction to kinetic Monte Carlo simulations of surface reactions*. Springer, 2012.

[95] Patanachai Janthon, Sergey M Kozlov, Francesc Vines, Jumras Limtrakul, and Francesc Illas. Establishing the accuracy of broadly used density functionals in describing bulk properties of transition metals. *Journal of Chemical Theory and Computation*, 9(3):1631–1640, 2013.

[96] Tao Jiang, D J Mowbray, Sergey Dobrin, H Falsig, Britt Hvolbæk, Thomas Bligaard, and Jens Kehlet Nørskov. Trends in CO oxidation rates for metal nanoparticles and close-packed, stepped, and kinked surfaces. *The Journal of Physical Chemistry C*, 113(24):10548–10553, 2009.

[97] Yaochu Jin, Senior Member, Markus Olhofer, and Bernhard Sendhoff. A Framework for Evolutionary Optimization With Approximate Fitness Functions. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 6(5):481–494, 2002.

[98] Ryosuke Jinnouchi, Kensaku Kodama Takahisa Suzuki, and Yu Morimoto. DFT calculations on electro-oxidations and dissolutions of Pt and PtAu nanoparticles. *Catalysis Today*, 262:100–109, 2016.

[99] Christos M. Kalamaras, George G. Olympiou, and Angelos M. Efstathiou. The water-gas shift reaction on Pt/gamma-Al2O3 catalyst: Operando SSITKA-DRIFTS-mass spectroscopy studies. *Catalysis Today*, 138(3-4):228–234, 2008.

[100] Evangelia Kalligiannaki, Markos A. Katsoulakis, and Petr Plechac. Spatial Two-Level Interacting Particle Simulations and Information Theory–Based Error Quantification. *SIAM Journal on Scientific Computing*, 36(2):A634–A667, 2014.

[101] Ayman M Karim, Vinay Prasad, Giannis Mpourmpakis, William W Lonergan, Anatoly I Frenkel, Jingguang G Chen, and Dionisios G Vlachos. Correlating Particle Size and Shape of Supported Ru / γ-Al 2 O 3 Catalysts with NH 3 Decomposition Activity. *Journal of the American Chemical Society*, 131(12):12230–12239, 2009.

[102] Markos A Katsoulakis and Petr Plechac. Information-theoretic tools for parametrized coarse-graining of non-equilibrium extended systems. *The Journal of Chemical Physics*, 139(7):74115, 2013.

[103] D Kim, B J Debusschere, and H N Najm. Spectral Methods for Parametric Sensitivity in Stochastic Dynamical Systems. *Biophysical Journal*, 92(2):379–393, 2007.

[104] Kwang-Ki K Kim and Richard D Braatz. Generalised polynomial chaos expansion approaches to approximate stochastic model predictive controlâ. *International Journal of Control*, 86(8):1324–1337, 2013.

[105] Kwang-Ki K Kim, Dongying Erin Shen, Zoltan K Nagy, and Richard D Braatz. Wiener's Polynomial Chaos for the Analysis and Control of Nonlinear Dynamical Systems with Probabilistic Uncertainties. *Control Systems, IEEE*, 33(October):58–67, 2013.

[106] Seok Ki Kim, Cheonghee Kim, Ji Hoon Lee, Jaeyoung Kim, Hyunjoo Lee, and Sang Heup Moon. Performance of shape-controlled Pd nanoparticles in the selective hydrogenation of acetylene. *Journal of Catalysis*, 306:146–154, 2013.

[107] John Kitchin. Preface: Trends in Computational Catalysis. *Topics in Catalysis*, 55(5-6):227–228, jun 2012.

[108] John R Kitchin, Neetha A Khan, Mark A Barteau, Jingguang G Chen, Boris Yakshinskiy, and Theodore E Madey. Elucidation of the active surface and origin of the weak metalhydrogen bond on Ni/Pt(1 1 1) bimetallic surfaces: a surface science and density functional theory study. *Surface Science*, 544(2-3):295–308, oct 2003.

[109] Charles Kittel. *Introduction to solid state physics*. John Wiley & Sons, Inc, New York, 2005.

[110] W. Kohn. Nobel Lecture: Electronic structure of matterwave functions and density functionals. *Reviews of Modern Physics*, 71(5):1253–1266, oct 1999.

[111] Marc T M Koper. Structure sensitivity and nanoscale effects in electrocatalysis. *Nanoscale*, 3(5):2054–2073, 2011.

[112] G Kresse and J Hafner. Ab initio molecular-dynamics simulation of the liquid-metalamorphous-semiconductor transition in germanium. *Physical Review B*, 49(20):14251, 1994.

[113] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.

[114] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169, 1996.

[115] Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558, 1993.

[116] Georg Kresse and D Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59(3):1758, 1999.

[117] Ilkeun Lee, Françoise Delbecq, Ricardo Morales, Manuel a Albiter, and Francisco Zaera. Tuning selectivity in catalysis by controlling particle shape. *Nature materials*, 8(2):132–138, 2009.

[118] Seung Woo Lee, Shuo Chen, Jin Suntivich, Kotaro Sasaki, Radoslav R Adzic, and Yang Shao-Horn. Role of surface steps of Pt nanoparticles on the electrochemical activity for oxygen reduction. *The Journal of Physical Chemistry Letters*, 1(9):1316–1320, 2010.

[119] Kurt Lejaeghere, Gustav Bihlmayer, T. Bjorkman, Peter Blaha, S. Blugel, Volker Blum, Damien Caliste, Ivano E Castelli, Stewart J Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, M. Du ak, Olle Eriksson, José A Flores-Livas, Kevin F Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, O. Granas, E K U Gross, Andris Gulans, François Gygi, D R Hamann, Phil J Hasnip, N A W Holzwarth, D. Iu an, Dominik B Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, E. Kucukbenli, Yaroslav O Kvashnin, Inka L M Locht, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, L. Nordstrom, Taisuke Ozaki, Lorenzo Paulatto, Chris J Pickard, Ward Poelmans, Matt I J Probert, Keith Refson, Manuel Richter, G.-M. Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, P. Thunstrom, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J van Setten, Veronique Van Speybroeck, John M Wills, Jonathan R Yates, G.-X. Zhang, and Stefaan Cottenier. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280):1415, 2016.

[120] Kurt Lejaeghere, Veronique Van Speybroeck, Guido Van Oost, and Stefaan Cottenier. Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals. *Critical Reviews in Solid State and Materials Sciences*, 39(1):1–24, 2014.

[121] Christopher Lester, Christian A. Yates, and Ruth E. Baker. Efficient parameter sensitivity computation for spatially-extended reaction networks. *The Journal of Chemical Physics*, 044106:1–14, 2016.

[122] H Li, M Zhao, and Q Jiang. Cohesive energy of clusters referenced by Wulff construction. *The Journal of Physical Chemistry C*, 113(18):7594–7597, 2009.

[123] Sen Lin, Jianyi Ma, Linsen Zhou, Caijin Huang, Daiqian Xie, and Hua Guo. Influence of Step Defects on Methanol Decomposition : Periodic Density Functional Studies on Pd (211) and Kinetic Monte Carlo Simulations. *Physical Chemistry C*, 117(211):451–459, 2013.

[124] Bård Lindström and Lars J. Pettersson. A Brief History of Catalysis. *CATTECH*, 7(4):130–138, 2003.

[125] Jeffery Ludwig and Dionisios G Vlachos. Ab initio molecular dynamics of hydrogen dissociation on metal surfaces using neural networks and novelty sampling. *The Journal of chemical physics*, 127(15):154716, 2007.

[126] Xianfeng Ma. Orbitalwise Coordination Number for Predicting Adsorption Properties of Metal Nanocatalysts. *Physical Review Letters*, 118(3):036101–036105, 2017.

[127] Chris A. Mack. Fifty Years of Moore's Law. *IEEE Transactions on Semiconductor Manufacturing*, 24(2):202–207, may 2011.

[128] Matteo Maestri, Dionisios G. Vlachos, Alessandra Beretta, Gianpiero Groppi, and Enrico Tronconi. Steam and dry reforming of methane on Rh: Microkinetic analysis and hierarchy of kinetic models. *Journal of Catalysis*, 259(2):211–222, 2008.

[129] Prasad S. Mahajan and Ricki G. Ingalls. Evaluation of methods used to detect warm-up period in steady state simulation. In R G Ingalls, MD Rossetti, JS Smith, and BA Peters, editors, *Proceedings of the 2004 Winter Simulation Conference*, pages 663–671, Piscataway, NJ, 2004. IEEE.

[130] Matthew D Marcinkowski, April D Jewell, Michail Stamatakis, Matthew B Boucher, Emily A Lewis, Colin J Murphy, Georgios Kyriakou, and E Charles H Sykes. Controlling a spillover pathway with the molecular cork effect. *Nature materials*, 12(6):523–528, 2013.

[131] Nenad Markovic, Hubert Gasteiger, and Philip N Ross. Kinetics of oxygen reduction on Pt (hkl) electrodes: implications for the crystallite size effect with supported Pt electrocatalysts. *Journal of the Electrochemical Society*, 144(5):1591–1597, 1997.

[132] R Timothy Marler and Jasbir S Arora. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395, 2004.

[133] Jacob A McGill, Babatunde A Ogunnaike, and Dionisios G Vlachos. Efficient gradient estimation using finite differencing and likelihood ratios for kinetic Monte Carlo simulations. *Journal of Computational Physics*, 231(21):7170–7186, 2012.

[134] Andrew J Medford, Jess Wellendorff, Aleksandra Vojvodic, Felix Studt, Frank Abild-Pedersen, Karsten W. Jacobsen, Thomas Bligaard, and J. K. Norskov. Assessing the reliability of calculated catalytic ammonia synthesis rates. *Science*, 345(6193):197–200, 2014.

[135] Prateek Mehta, Jeffrey Greeley, W. Nicholas Delgass, and William F. Schneider. Adsorption Energy Correlations at the Metal-Support Boundary. *ACS Catalysis*, 7(7):4707–4715, jun 2017.

[136] L. R. Merte, G. Peng, R. Bechstein, F. Rieboldt, C. A. Farberow, L. C. Grabow, W. Kudernatsch, S. Wendt, E. Laegsgaard, M. Mavrikakis, and F. Besenbacher. Water-Mediated Proton Hopping on an Iron Oxide Surface. *Science*, 336(6083):889–893, may 2012.

[137] Hakim Meskine, Sebastian Matera, Matthias Scheffler, Karsten Reuter, and Horia Metiu. Examination of the concept of degree of rate control by first-principles kinetic Monte Carlo simulations. *Surface Science*, 603:1724–1730, 2009.

[138] A B Mhadeshwar, H Wang, and D G Vlachos. Thermodynamic Consistency in Microkinetic Development of Surface Reaction Mechanisms. *The Journal of Physical Chemistry B*, 107(46):12721–12733, 2003.

[139] A.B. Mhadeshwar, J.R. Kitchin, M.A. Barteau, and D.G. Vlachos. The Role of Adsorbate-Adsorbate Interactions in the Rate Controlling Step and the Most Abundant Reaction Intermediate of NH3 Decomposition on Ru. *Catalysis Letters*, 96(1/2):13–22, jul 2004.

[140] A.B. Mhadeshwar and D.G. Vlachos. Is the water-gas shift reaction on Pt simple?: Computer-aided microkinetic model reduction, lumped rate expression, and rate-determining step. *Catalysis Today*, 105(1):162–172, 2005.

[141] Spencer D. Miller and John R. Kitchin. Uncertainty and figure selection for DFT based cluster expansions for oxygen adsorption on Au and Pt (111) surfaces. *Molecular Simulation*, 35(10-11):920–927, 2009.

[142] Spencer D. Miller, Nilay nolu, and John R. Kitchin. Configurational correlations in the coverage dependent adsorption energies of oxygen atoms on late transition metal fcc(111) surfaces. *The Journal of Chemical Physics*, 134:104709, mar 2011.

[143] Alexander V. Mironenko and Dionisios G. Vlachos. Conjugation-Driven "reverse Mars-van Krevelen"-Type Radical Mechanism for Low-Temperature C-O Bond Activation. *Journal of the American Chemical Society*, 138(26):8104–8113, 2016.

[144] Simon Mostafa, Farzad Behafarid, Jason R Croy, Luis K Ono, Long Li, Judith C Yang, Anatoly I Frenkel, and Beatriz Roldan Cuenya. Shape-dependent catalytic properties of Pt nanoparticles. *Journal of the American Chemical Society*, 132(44):15714–15719, 2010.

[145] Giannis Mpourmpakis, Antonis N Andriotis, and Dionisios G Vlachos. Identification of descriptors for the CO interaction with metal nanoparticles. *Nano letters*, 10(3):1041–1045, 2010.

[146] E. P. Munger and M. A. Novotny. Reweighting in Monte Carlo and Monte Carlo renormalization-group studies. *Physical Review B*, 43(7):5773–5783, 1991.

[147] Marvin K Nakayama, Ambuj Goyal, and Peter W Glynn. Likelihood ratio sensitivity analysis for Markovian models of highly dependable systems. *Operations Research*, 42(1):137–157, 1994.

[148] Jens Nielsen, Mayeul D'Avezac, James Hetherington, and Michail Stamatakis. Parallel kinetic Monte Carlo simulation framework incorporating accurate models of adsorbate lateral interactions. *Journal of Chemical Physics*, 139(22):224706, 2013.

[149] Nima Nikbin, Natalie Austin, Dionisios G. Vlachos, Michail Stamatakis, and Giannis Mpourmpakis. Catalysis at the sub-nanoscale: complex CO oxidation chemistry on a few Au atoms. *Catal. Sci. Technol.*, 5(1):134–141, dec 2015.

[150] J K Norskov, T Bligaard, J Rossmeisl, and C H Christensen. Towards the computational design of solid catalysts. *Nat Chem*, 1(1):37–46, 2009.

[151] Jens K Nørskov, Thomas Bligaard, Britt Hvolbaek, Frank Abild-Pedersen, Ib Chorkendorff, Claus H Christensen, J K Norskov, Thomas Bligaard, Britt Hvolbaek, Frank Abild-Pedersen, Ib Chorkendorff, and Claus H Christensen. The nature of the active site in heterogeneous metal catalysis. *Chemical Society Reviews*, 37(10):2163–2171, 2008.

[152] Jens Kehlet Nørskov, Jan Rossmeisl, Ashildur Logadottir, Lars Lindqvist, John R Kitchin, Thomas Bligaard, and Hannes Jonsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, 2004.

[153] Yaghout Nourani and Bjarne Andresen. A comparison of simulated annealing cooling strategies. *Journal of Physics A: Mathematical and General*, 31(41):8373, 1998.

[154] M. Núñez, T. Robie, and D. G. Vlachos. Acceleration and sensitivity analysis of lattice kinetic Monte Carlo simulations using parallel processing and rate constant rescaling. *The Journal of Chemical Physics*, 147(16):164103, oct 2017.

[155] M Núñez and D G Vlachos. Steady state likelihood ratio sensitivity analysis for stiff kinetic Monte Carlo simulations. *The Journal of Chemical Physics*, 142(4):44108, 2015.

[156] Gen Ouyang, Kong-Jie Zhu, Lei Zhang, Peng-Fei Cui, Bo-Tao Teng, and Xiao-Dong Wen. Effects of coordination number of Au catalyst on oxygen species and their catalytic roles. *Applied Surface Science*, 387:875–881, nov 2016.

[157] Stanislav Pandelov and Ulrich Stimming. Reactivity of monolayers and nano-islands of palladium on Au(1 1 1) with respect to proton reduction. *Electrochimica Acta*, 52(18):5548–5555, may 2007.

[158] Yannis Pantazis and Markos A. Katsoulakis. A relative entropy rate method for path space sensitivity analysis of stationary complex stochastic dynamics. *The Journal of Chemical Physics*, 138(5):54115, 2013.

[159] Yannis Pantazis, Markos A Katsoulakis, and Dionisios G Vlachos. Parametric sensitivity analysis for biochemical reaction networks based on pathwise information theory. *BMC bioinformatics*, 14(1):311, 2013.

[160] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

[161] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865, 1996.

[162] Francisco J. Perez-Alonso, David N. McCarthy, Anders Nierhoff, Patricia Hernandez-Fernandez, Christian Strebel, Ifan E L Stephens, Jane H. Nielsen, and Ib Chorkendorff. The effect of size on the oxygen electroreduction activity of mass-selected platinum nanoparticles. *Angewandte Chemie - International Edition*, 51(19):4641–4643, 2012.

[163] Linda Petzold and Wenjie Zhu. Model reduction for chemical kinetics: An optimization approach. *American Institute of Chemical Engineers. AIChE Journal*, 45(4):869, 1999.

[164] Simone Piccinin and Michail Stamatakis. CO Oxidation on Pd(111): A First-Principles-Based Kinetic Monte Carlo Study. *ACS Catalysis*, 4(7):2143–2152, 2014.

[165] Simone Piccinin and Michail Stamatakis. Steady-State CO Oxidation on Pd(111): First-Principles Kinetic Monte Carlo Simulations and Microkinetic Analysis. *Topics in Catalysis*, 60(1-2):141–151, feb 2017.

[166] Evgeny A Pidko. Towards the balance between the reductionist and systems approaches in computational catalysis: model versus method accuracy for the description of catalytic systems. *ACS Catalysis*, 7:4230–4234, 2017.

[167] Sergey Plyasunov and Adam P Arkin. Efficient stochastic sensitivity analysis of discrete event systems. *Journal of Computational Physics*, 221(2):724–738, 2007.

[168] H Rabitz, M Kramer, D Dacol, Annual Reviews, H Rabitz, M Kramer, and D Dacol. Sensitivity analysis in chemical kinetics. *Annual Review of Physical Chemistry*, 34(1):419–461, 1983.

[169] S Raimondeau and D G Vlachos. The role of adsorbate-layer nonuniformities in catalytic reactor design: multiscale simulations for CO oxidation on Pt. *Computers & Chemical Engineering*, 26(7-8):965–980, 2002.

[170] Christopher V. Rao and Adam P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *Journal of Chemical Physics*, 118(11):4999–5010, 2003.

[171] M Rathinam, L R Petzold, Y Cao, and D T Gillespie. Stiffness in stochastically reacting systems: The implicit tau-leaping method. *Journal of Chemical Physics*, 119(24):12784–12794, 2003.

[172] Muruhan Rathinam, Patrick W. Sheppard, and Mustafa Khammash. Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *Journal of Chemical Physics*, 132(3):034103, 2010.

[173] James Russell Renzas, Yawen Zhang, Wenyu Huang, and Gabor A. Somorjai. Rhodium nanoparticle shape dependence in the reduction of NO by CO. *Catalysis Letters*, 132(3-4):317–322, 2009.

[174] Haluk Resat, Linda Petzold, and Michel F Pettigrew. Kinetic modeling of biological systems. In *Computational Systems Biology*, pages 311–335. Springer, 2009.

[175] Karsten Reuter. First-Principles Kinetic Monte Carlo Simulations for Heterogeneous Catalysis: Concepts , Status and Frontiers. *Modeling Heterogeneous Catalytic Reactions: From the Molecular Process to the Technical System*, 2011.

[176] Karsten Reuter, Craig P. Plaisance, Harald Oberhofer, and Mie Andersen. Perspective: On the active site model in computational catalyst screening. *The Journal of Chemical Physics*, 146(4):040901, 2017.

[177] B. Roldan Cuenya, A. I. Frenkel, S. Mostafa, F. Behafarid, J. R. Croy, L. K. Ono, and Q. Wang. Anomalous lattice dynamics and thermal properties of supported size- and shape-selected Pt nanoparticles. *Physical Review B*, 82(15):155450, oct 2010.

[178] Beatriz Roldan Cuenya. Metal nanoparticle catalysts beginning to shape-up. *Accounts of Chemical Research*, 46(8):1682–1691, 2012.

[179] Xi Rong, Jules Parolin, and Alexie M Kolpak. A fundamental relationship between reaction mechanism and stability in metal oxide catalysts for oxygen evolution. *ACS Catalysis*, 6(2):1153–1158, 2016.

[180] Jan Rossmeisl, Gustav S Karlberg, Thomas Jaramillo, and Jens K Nørskov. Steady state oxygen reduction and cyclic voltammetry. *Faraday Discussions*, 140:337–346, 2009.

[181] Jakob Ruess, Heinz Koeppl, and Christoph Zechner. Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability. *The Journal of Chemical Physics*, 146(12):124122, mar 2017.

[182] M. Salciccioli, Y. Chen, and D. G. Vlachos. Density functional theory-derived group additivity and linear scaling methods for prediction of oxygenate stability on metal catalysts: Adsorption of open-ring alcohol and polyol dehydrogenation intermediates on pt-based metals. *Journal of Physical Chemistry C*, 114(47):20155–20166, 2010.

[183] M Salciccioli, M Stamatakis, S Caratzoulas, and D G Vlachos. A review of multiscale modeling of metal-catalyzed reactions: Mechanism development for complexity and emergent behavior. *Chemical Engineering Science*, 66(19):4319–4355, 2011.

[184] Michael Salciccioli and D G Vlachos. Kinetic modeling of pt catalyzed and computation-driven catalyst discovery for ethylene glycol decomposition. *ACS Catalysis*, 1(10):1246–1256, 2011.

[185] Howard Salis and Yiannis Kaznessis. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *Journal of Chemical Physics*, 122(5):054103, 2005.

[186] Howard Salis, Vassilios Sotiropoulos, and Yiannis N Kaznessis. Multiscale Hy3S: Hybrid stochastic simulation for supercomputers. *BMC bioinformatics*, 7(1):93, 2006.

[187] A. Samant and D. G. Vlachos. Overcoming stiffness in stochastic simulation stemming from partial equilibrium: A multiscale Monte Carlo algorithm. *Journal of Chemical Physics*, 123(14):144114, 2005.

[188] Asawari Samant, Babatunde A Ogunnaike, and Dionisios G Vlachos. A hybrid multiscale Monte Carlo algorithm (HyMSMC) to cope with disparity in time scales and species populations in intracellular networks. *BMC bioinformatics*, 8:175–197, 2007.

[189] Sergio I. Sanchez, Laurent D. Menard, Ariella Bram, Joo H. Kang, Matthew W. Small, Ralph G. Nuzzo, and Anatoly I. Frenkel. The Emergence of Nonbulk Properties in Supported Metal Clusters: Negative Thermal Expansion and Atomic Disorder in Pt Nanoclusters Supported on γ-Al 2 O 3. *Journal of the American Chemical Society*, 131(20):7040–7054, may 2009.

[190] David J. Schmidt, Wei Chen, C. Wolverton, and William F. Schneider. Performance of Cluster Expansions of Coverage-Dependent Adsorption of Atomic Oxygen on Pt (111). *Journal of Chemical Theory and Computation*, 8(1):264–273, 2012.

[191] Sanjaya D. Senanayake, José A. Rodriguez, and Dario Stacchiola. Electronic MetalSupport Interactions and the Production of Hydrogen Through the Water-Gas Shift Reaction and Ethanol Steam Reforming: Fundamental Studies with Well-Defined Model Catalysts. *Topics in Catalysis*, 56(15-17):1488–1498, 2013.

[192] Richard Serfozo. *Basics of Applied Stochastic Processes*. Probability and Its Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[193] David A. Sheen, Xiaoqing You, Hai Wang, and Terese Lovas. Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. *Proceedings of the Combustion Institute*, 32(1):535–542, 2009.

[194] Jeffrey J. Siirola. The impact of shale gas in the chemical industry. *AIChE Journal*, 60(3):810–819, mar 2014.

[195] M.A. Snyder, A. Chatterjee, and D.G. Vlachos. Net-event kinetic Monte Carlo for overcoming stiffness in spatially homogeneous and distributed systems. *Computers & Chemical Engineering*, 29(4):701–712, 2005.

[196] Gabor A. Somorjai and Jeong Y. Park. Colloid Science of Metal Nanoparticle Catalysts in 2D and 3D Structures. Challenges of Nucleation, Growth, Composition, Particle Shape, Size Control and Their Influence on Activity and Selectivity. *Topics in Catalysis*, 49(3-4):126–135, 2008.

[197] Sergio Filipe Sousa, Pedro Alexandrino Fernandes, and Maria Joao Ramos. General performance of density functionals. *Journal of Physical Chemistry A*, 111(42):10439–10452, 2007.

[198] Fabian Spill, Philip K. Maini, and Helen M. Byrne. Optimisation of simulations of stochastic processes by removal of opposing reactions. *Journal of Chemical Physics*, 144(8):84105, 2016.

[199] Rishi Srivastava, David F. Anderson, and James B. Rawlings. Comparison of finite difference based methods to obtain sensitivities of stochastic chemical kinetic models. *Journal of Chemical Physics*, 138(7):074110, 2013.

[200] M. Stamatakis. Kinetic modelling of heterogeneous catalytic systems. *J Phys Condens Matter*, 27(1):13001–13028, 2015.

[201] Michail Stamatakis, Ying Chen, and Dionisios G Vlachos. First-Principles-Based Kinetic Monte Carlo Simulation of the Structure Sensitivity of the Water Gas Shift Reaction on Platinum Surfaces. *The Journal of Physical Chemistry C*, 115(50):24750–24762, 2011.

[202] Michail Stamatakis, Matthew A Christiansen, Dionisios G Vlachos, and Giannis Mpourmpakis. Multiscale modeling reveals poisoning mechanisms of MgO-supported Au clusters in CO oxidation. *Nano letters*, 12(7):3621–3626, 2012.

[203] Michail Stamatakis and Simone Piccinin. Rationalizing the Relation between Adlayer Structure and Observed Kinetics in Catalysis. *ACS Catalysis*, 6(3):2105–2111, mar 2016.

[204] Michail Stamatakis and Dionisios G. Vlachos. A graph-theoretical kinetic Monte Carlo framework for on-lattice chemical kinetics. *The Journal of Chemical Physics*, 134(21):214115–214127, 2011.

[205] Michail Stamatakis and Dionisios G. Vlachos. Equivalence of on-lattice stochastic chemical kinetics with the well-mixed chemical master equation in the limit of fast diffusion. *Computers and Chemical Engineering*, 35(12):2602–2610, 2011.

[206] Michail Stamatakis and Dionisios G Vlachos. Unraveling the Complexity of Catalytic Reactions via Kinetic Monte Carlo Simulation: Current Status and Frontiers. *ACS Catalysis*, 2(12):2648–2663, 2012.

[207] Carsten Stegelmann, Anders Andreasen, and Charles T Campbell. Degree of rate control: How much the energies of intermediates and transition states control rates. *Journal of the American Chemical Society*, 131(23):8077–8082, 2009.

[208] Ifan E. L. Stephens, Alexander S. Bondarenko, Ulrik Grønbjerg, Jan Rossmeisl, and Ib Chorkendorff. Understanding the electrocatalysis of oxygen reduction on platinum and its alloys. *Energy & Environmental Science*, 5(5):6744, 2012.

[209] Jianwei Sun, Adrienn Ruzsinszky, and JohnP. Perdew. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters*, 115(3):036402, jul 2015.

[210] Jonathan E. Sutton, Wei Guo, Markos A. Katsoulakis, and Dionisios G. Vlachos. Effects of correlated parameters and uncertainty in electronic-structure-based chemical kinetic modelling. *Nature Chemistry*, 8(4):331–337, 2016.

[211] Jonathan E. Sutton and Dionisios G. Vlachos. Error estimates in semi-empirical estimation methods of surface reactions. *Journal of Catalysis*, 297:202–216, 2013.

[212] Jonathan E. Sutton and Dionisios G. Vlachos. Effect of errors in linear scaling relations and Brønsted-Evans-Polanyi relations on activity and selectivity maps. *Journal of Catalysis*, 338:273–283, 2016.

[213] Burcin Temel, Hakim Meskine, Karsten Reuter, Matthias Scheffler, and Horia Metiu. Does phenomenological kinetics provide an adequate description of heterogeneous catalytic reactions? *Journal of Chemical Physics*, 126(20):204711, 2007.

[214] D. Tomanek, S. Mukherjee, and K. H. Bennemann. Simple theory for the electronic and atomic structure of small clusters. *Phys. Rev. B*, 28(2):665–673, 1983.

[215] Vladimir Tripković, Isotta Cerri, Thomas Bligaard, and Jan Rossmeisl. The Influence of Particle Shape and Size on the Activity of Platinum Nanoparticles for Oxygen Reduction Reaction: A Density Functional Theory Study. *Catalysis Letters*, 144(3):380–388, 2014.

[216] G A Tritsaris, J Greeley, Jan Rossmeisl, and Jens Kehlet Nørskov. Atomic-scale modeling of particle size effects for the oxygen reduction reaction on Pt. *Catalysis Letters*, 141(7):909–913, 2011.

[217] Sarah A. Tupy, Ayman M. Karim, Christina Bagia, Weihua Deng, Yulin Huang, Dionisios G. Vlachos, and Jingguang G. Chen. Correlating ethylene glycol reforming activity with in situ EXAFS detection of Ni segregation in supported NiPt bimetallic catalysts. *ACS Catalysis*, 2(11):2290–2296, 2012.

[218] W R Tyson and W A Miller. Surface free energies of solid metals: Estimation from liquid surface tension measurements. *Surface Science*, 62(1):267–276, 1977.

[219] Z Ulissi, V Prasad, and D G Vlachos. Effect of multiscale model uncertainty on identification of optimal catalyst properties. *Journal of Catalysis*, 281(2):339–344, 2011.

[220] Zachary W. Ulissi, Andrew J. Medford, Thomas Bligaard, Jens K. Nørskov, and J. K. Nørskov. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nature Communications*, 8:14621, 2017.

[221] Zachary W Ulissi, Michael T Tang, Jianping Xiao, Xinyan Liu, Daniel A Torelli, Mohammadreza Karamad, Kyle Cummins, Christopher Hahn, Nathan S Lewis, Thomas F Jaramillo, Karen Chan, and Jens K Nørskov. Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO2 Reduction. *ACS Catalysis*, 7(10):6600–6608, 2017.

[222] Mauro Valorani, Francesco Creta, Dimitris A. Goussis, Jeremiah C. Lee, and Habib N. Najm. An automatic procedure for the simplification of chemical kinetic mechanisms based on CSP. *Combustion and Flame*, 146(1-2):29–51, 2006.

[223] Rutger A. Van Santen. Complementary Structure Sensitive and Insensitive Catalytic Relationships. *Accounts of Chemical Research*, 42(1):57–66, 2009.

[224] Dionisios G Vlachos. A Review of Multiscale Analysis: Examples from Systems Biology, Materials Engineering, and Other FluidSurface Interacting Systems. *Advances in Chemical Engineering*, 30:1–61, 2005.

[225] Aleksandra Vojvodic, Andrew James Medford, Felix Studt, Frank Abild-Pedersen, Tuhin Suvra Khan, T Bligaard, and J K Nørskov. Exploring the limits: A low-pressure, low-temperature HaberBosch process. *Chemical Physics Letters*, 598:108–112, 2014.

[226] Bin Wang. Recent development of non-platinum catalysts for oxygen reduction reaction. *Journal of Power Sources*, 152:1–15, 2005.

[227] Hangyao Y Wang, Michail Stamatakis, Danielle A. Hansgen, Stavros Caratzoulas, and Dionisios G. Vlachos. Understanding mixing of Ni and Pt in the Ni/Pt(111) bimetallic catalyst via molecular simulation and experiments. *Journal of Chemical Physics*, 133(22):224503–224513, 2010.

[228] Liya Wang, Ulrich Stimming, and Michael Eikerling. Kinetic Model of Hydrogen Evolution at an Array of Au-Supported Catalyst Nanoparticles. *Electrocatalysis*, 1(1):60–71, may 2010.

[229] S. Wang, V. Petzold, V. Tripkovic, J. Kleis, J. G. Howalt, E. Skúlason, E. M. Fernández, B. Hvolbæk, G. Jones, a. Toftelund, H. Falsig, M. Björketun, F. Studt, F. Abild-Pedersen, J. Rossmeisl, J. K. Nørskov, and T. Bligaard. Universal transition state scaling relations for (de)hydrogenation over transition metals. *Physical Chemistry Chemical Physics*, 13(46):20760, 2011.

[230] Patrick B Warren and Rosalind J Allen. Steady-state parameter sensitivity in stochastic modeling via trajectory reweighting. *The Journal of Chemical Physics*, 136(10):104106, 2012.

[231] Jess Wellendorff, Trent L Silbaugh, Delfina Garcia-Pintos, Jens K Nørskov, Thomas Bligaard, Felix Studt, and Charles T Campbell. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals. *Surface Science*, 640:36–44, 2015.

[232] Jaime Wisniak. The history of catalysis. From the beginning to Nobel Prizes. *Educ. quim.*, 21(1):60–69, 2010.

[233] Christopher A. Wolcott, Andrew J. Medford, Felix Studt, and Charles T. Campbell. Degree of rate control approach to computational catalyst screening. *Journal of Catalysis*, 330:197–207, 2015.

[234] C Wu, D J Schmidt, C Wolverton, and W F Schneider. Accurate coverage-dependence incorporated into first-principles kinetic models: Catalytic NO oxidation on Pt (111). *Journal of Catalysis*, 286:88–94, 2012.

[235] Xiaohu Xia, Sang-Il Choi, Jeffrey A Herron, Ning Lu, Jessica Scaranto, Hsin-Chieh Peng, Jinguo Wang, Manos Mavrikakis, Moon J Kim, and Younan Xia. Facile synthesis of palladium right bipyramids and their use as seeds for overgrowth and as catalysts for formic acid oxidation. *Journal of the American Chemical Society*, 135(42):15706–15709, 2013.

[236] Kimihisa Yamamoto, Takane Imaoka, Wang-Jae Chun, Osamu Enoki, Hideaki Katoh, Masahiro Takenaga, and Atsunori Sonoi. Size-specific catalytic activity of platinum clusters enhances oxygen reduction reactions. *Nature chemistry*, 1(5):397–402, 2009.

[237] Liu Yang, Altaf Karim, and James T Muckerman. Density Functional Kinetic Monte Carlo Simulation of WaterGas Shift Reaction on Cu/ZnO. *The Journal of Physical Chemistry C*, 117(7):3414–3425, 2013.

[238] George Yin and Qing Zhang. *Continuous-time Markov chains and applications: a two-time-scale approach*, volume 37. Springer Science & Business Media, second edition, 2012.

[239] Yongsheng Zhang, Volker Blum, and Karsten Reuter. Accuracy of first-principles lateral interactions: Oxygen at Pd (100). *Physical Review B*, 75(23):235406, 2007.

[240] Houlong Zhuang, Alexander J Tkalych, and Emily A Carter. Surface Energy as a Descriptor of Catalytic Activity. *The Journal of Physical Chemistry C*, 120(41):23698–23706, 2016.

Appendix A

EFFICIENT SENSITIVITY ANALYSIS IN MULTISCALE KINETIC MONTE CARLO - SUPPLEMENT

A.1 Sensitivity analysis in multiscale deterministic systems

In this section, we look at the time-evolution of a two time-scale kinetic problem. It is solved in two ways. First, a stiff, single time-scale (STS) formulation is used. Secondly, a two time-scale (TTS) formulation is used. Methods for performing on-line sensitivity analysis are demonstrated for each. The TTS formalism converts the stiff system of ODEs into a system of differential algebraic equations.

A.1.1 Single time scale (STS) Formulation

In a well mixed chemical system, the vector N of species populations completely specifies the system. The rate of each of the elementary reactions is represented by the column vector $r(N)$ as a function of N. $r(N)$ can have a non-linear dependence on N when the network contains multimolecular reactions. The stoichiometric matrix S allows us to compute differential changes in species populations due to the elementary reactions. The changes in species populations can be written as a system of ordinary differential equations (ODE) which can be integrated from time $t = 0$ to $t = t_{\text{final}}$ to obtain $N(t_{\text{final}})$.

$$N(0) = N_0 \tag{A.1}$$

$$\frac{dN}{dt} = S \cdot r(N) \tag{A.2}$$

Consider that the calculation of $r(N)$ involves model parameters k_α and therefore $r = r(N, k_\alpha)$. k_α is typically either a rate constant or a thermodynamic parameter

used to compute the rate constants. The resulting time dependent species populations $N(t)$ also depend on the parameters k_α. For each parameter k_α of interest, we are interested in the sensitivity coefficient $C_{k_\alpha}(t) = \frac{\partial N(t)}{\partial k_\alpha}$. By applying the chain rule, derivatives of a system of ODEs are computed by solving an adjunct set of ODEs which are integrated alongside the original system. Taking the derivatives of Equations A.1 and A.2 with respect to k_α results in Equations A.3 and A.4 respectively.

$$C_{k_\alpha}(0) = \mathbf{0}_{M\times 1} \tag{A.3}$$

$$\frac{dC_{k_\alpha}}{dt} = S \cdot \left[\frac{\partial r(N,k_\alpha)}{\partial N} C_{k_\alpha} + \frac{\partial r(N,k_\alpha)}{\partial k_\alpha}\right] \tag{A.4}$$

The matrix $\frac{\partial r(N,k_\alpha)}{\partial N}$ and the vector $\frac{\partial r(N,k_\alpha)}{\partial k_\alpha}$ must be known analytical functions. However, usually the rate law $r(N,k_\alpha)$ is a simple enough polynomial function that this is not an issue.

A.1.2 Two time scale (TTS) Formulation

Separation of time scales is extremely common in kinetic systems. It is often advantageous to separate the system into two time scales. Assumptions such as partial equilibrium impose constraints on the fast time scale. Algebraic equations can be used to determine the fast-changing variables as a result. The solution to these equations can be fed into an ODE for the slowly evolving variables of the system. The TTS formulation is both computationally and pedagogically advantageous.

We would like to have our system in the form

$$0 = g(y_s, y_f) \tag{A.5}$$

$$\frac{dy_s}{dt} = f(y_f, y_s, \theta) \tag{A.6}$$

First a transformation of the system variables N to a set of variables $y = \begin{bmatrix} y_f \\ y_s \end{bmatrix}$ must take place. The variables y are decoupled into variables y_f which evolve on the fast scale and variables y_s which evolve only along the slow scale. Upon identifying

which reactions are fast, the stoichiometric matrix can be decomposed into fast and slow contributions.

$$S = S_s + S_f \tag{A.7}$$

We seek a transformation $T = \begin{bmatrix} T_f \\ T_s \end{bmatrix}$ on N to produce

$$y = T \cdot N \tag{A.8}$$

We use Gaussian elimination on S_f to find a transformation matrix T such that all of the linearly dependent rows are replaced with 0s. T is a transformation matrix that is a series of row operations and is therefore invertible. That is,

$$S'_f = T \cdot S_f \tag{A.9}$$

If a row i of S'_f contains all 0s, then the ith row of T belongs to T_s. These are slow variables $y_s = T_s \cdot N$ which do not change on the fast scale. Nonzero rows correspond to fast variables given by $y_f = T_f \cdot N$.

There are m_f fast modes and $m_s = M - m_f$ slow modes. On the fast scale we solve the system of equations for y_f given y_s, because y_f is slaved to y_s. We substitute $N = T^{-1} \cdot y$ to eliminate the functional dependence of r on N.

$$0 = T_f \cdot S_f \cdot r \left(T^{-1} \cdot \begin{bmatrix} y_f \\ y_s \end{bmatrix} \right) \tag{A.10}$$

Equation A.10 is a system of n_f equations which encompass the equilibrium constraint on the fast scale. On the slow scale we solve the system of n_s differential equations

$$y_s(0) = T_s \cdot N_0 \tag{A.11}$$

$$\frac{dy_s}{dt} = T_s \cdot S_s \cdot r \left(T^{-1} \cdot \begin{bmatrix} y_f \\ y_s \end{bmatrix} \right) \tag{A.12}$$

At every macroscopic ODE step, we solve the fast scale to get y_f, plug it into the differential equation of Equation A.12 to advance y_s, then use y_s to solve for the

new y_f and repeat until we have reached t_{final}. Upon solving the ODE, we can recover the original system variables using $N = T^{-1} \cdot y$.

The approach for sensitivity analysis is similar to that for STS, but derivatives must be computed from a system of equations involving the fast time scale. For each parameter k_α of interest, taking the derivatives of Equations A.10, A.11, and A.12 results in the following system for $C^y_{k_\alpha}(t) = \frac{\partial y(t)}{\partial k_\alpha}$, $C^y_{f,k_\alpha}(t) = \frac{\partial y_f(t)}{\partial k_\alpha}$, and $C^y_{s,k_\alpha}(t) = \frac{\partial y_s(t)}{\partial k_\alpha}$.

$$0 = T_f \cdot S_f \cdot \left(\frac{\partial r(N, k_\alpha)}{\partial N} \cdot \frac{\partial N}{\partial y} \cdot \begin{bmatrix} C^y_{f,k_\alpha} \\ C^y_{s,k_\alpha} \end{bmatrix} + \frac{\partial r(N, k_\alpha)}{\partial k_\alpha} \right) \tag{A.13}$$

$$C^y_{s,k_\alpha}(0) = \mathbf{0}_{m_s \times 1} \tag{A.14}$$

$$\frac{dC^y_{s,k_\alpha}}{dt} = T_s \cdot S_s \cdot \left(\frac{\partial r(N, k_\alpha)}{\partial N} \cdot \frac{\partial N}{\partial y} \cdot \begin{bmatrix} C^y_{f,k_\alpha} \\ C^y_{s,k_\alpha} \end{bmatrix} + \frac{\partial r(N, k_\alpha)}{\partial k_\alpha} \right) \tag{A.15}$$

At each step of the ODE, Equation A.13 solves for C^y_{f,k_α} given C^y_{s,k_α}. Then, C^y_{f,k_α} is used in Equation A.15 to compute the time evolution of C^y_{s,k_α}. In these equations, we use $\frac{dN}{dy} = T^{-1}$, which we can easily get from Equation A.8. At the end of the procedure, the derivatives of the original system variables can be obtained using $C_{k_\alpha} = T^{-1} \cdot C^y_{k_\alpha}$.

A.1.3 Numerical example

We use the same model system and parameters as in Section 2.5.3. The system contains $M = 3$ species and $R = 3$ reactions. Mathematically, we use the $M \times 1$ column vector N to specify the species populations, where $N_1 = N_A$, $N_2 = N_B$, and $N_3 = N_C$. In this example, there is only a linear dependence and the reaction rates can be written as

$$r(N) = \begin{bmatrix} k_1 N_1 \\ k_2 N_2 \\ k_3 N_2 \end{bmatrix}. \tag{A.16}$$

The stoichiometric matrices are

$$S = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \tag{A.17}$$

$$S_s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \tag{A.18}$$

$$S_f = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{A.19}$$

A suitable transformation is

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{A.20}$$

resulting in

$$S'_f = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{A.21}$$

T can be decomposed into $T_{\text{fast}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ and $T_{\text{slow}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ by looking at the 0 rows of S'_f. This gives us the transformed variables as

$$y_{\text{fast}} = \begin{bmatrix} N_1 \end{bmatrix} \tag{A.22}$$

$$y_{\text{slow}} = \begin{bmatrix} N_1 + N_2 \\ N_3 \end{bmatrix} \tag{A.23}$$

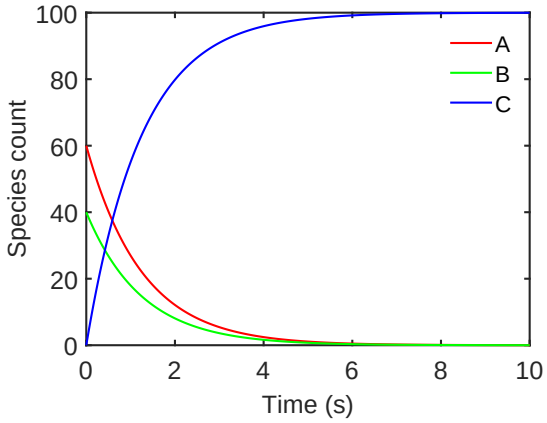
In the context of our example problem, we can assign physical meaning to the transformation. $y_1 = N_A$ is affected by both slow and fast reactions. For a given set of slow variables, we can solve for y_1 to specify the equilibrium constraint of $r_1 = r_2$.

$y_2 = N_A + N_B$ is the total amount of A and B, which is unaffected by the fast $A \leftrightarrow B$ reaction, but is affected by the slow $B \rightarrow C$ reaction. $y_3 = N_C$ is a second slow variable which is unaffected by fast reactions, but evolves on the slow manifold. The mode $N_A + N_B + N_C$ does not appear as a fast mode constraint. If Gaussian elimination were done on the full matrix S, this mode would be identified. Likewise, it could appear as a "slow mode" according to our labeling. It is not a "fast mode" in our terminology because the mass conservation is not a consequence of the constraints imposed by partial equilibrium, captured in the structure of S_f.

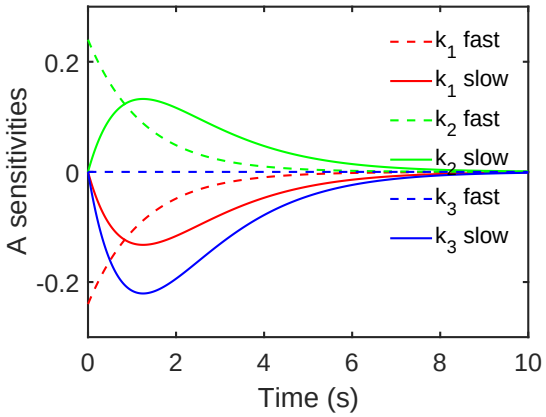
We use ODE15s in Matlab to solve the system of differential equations. For the TTS, a mass matrix is used to convert the system to a differential algebraic one. Results are shown in Figure A.1. Partitioning the sensitivities between the fast and slow contributions offers several interesting physical insights. For species B in Figure A.1c, the fast and slow contributions of k_1 have competing effects on the population of B. On the fast scale, higher k_1 leads to a higher equilibrium quantity of species B. However, on the slow scale, the greater amount of B relative to A causes faster loss of species B to C. Therefore, the sensitivity of N_B to k_1 is positive on the fast scale, but negative on the slow scale. In Figure A.1d, all sensitivity contributions to N_C are zero on the fast scale, as species C is only affected by slow reactions and evolves on the slow manifold only.

A.2 Analytical solution for the A↔B→C System

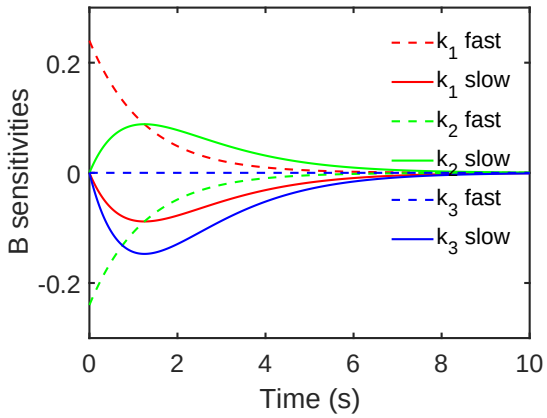
We have derived analytical expressions for the transient species populations and sensitivityies of our model system (see Section 2.5.3) in the two time scale limit. In all equations, $K = \frac{k_1}{k_2}$ is the equilibrium constant of the A↔B reaction.



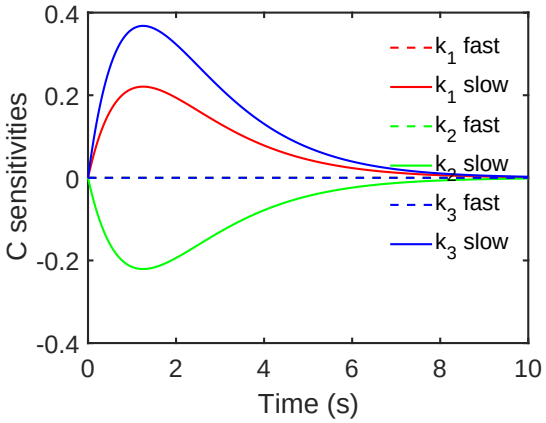
(a) Species populations versus time.



(b) N_A derivatives versus time.



(c) N_B derivatives versus time.



(d) N_C derivatives versus time.

Figure A.1: Results of two time scale simulation of the model chemical reaction network. Sensitivities are normalized as in Equation 2.32 and are partioned into contributions from the fast and slow time scales. Derivative can be obtained by adding the contributions from the two time scales.

$$A(t) = A(0) * \frac{1}{K+1}\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.24}$$

$$B(t) = A(0) * \frac{K}{K+1}\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.25}$$

$$C(t) = A(0) * \left(1 - \exp\left(-\frac{k_3 K t}{K+1}\right)\right) \tag{A.26}$$

$$\frac{\partial A(t)}{\partial k_1} = A(0)\left[\frac{-k_2}{(k_1+k_2)^2} + \left(\frac{1}{K+1}\right)\frac{k_2}{(k_1+k_2)^2}(-k_3 t)\right]\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.27}$$

$$\frac{\partial A(t)}{\partial k_2} = A(0)\left[\frac{k_1}{(k_1+k_2)^2} + \left(\frac{1}{K+1}\right)\frac{-k_1}{(k_1+k_2)^2}(-k_3 t)\right]\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.28}$$

$$\frac{\partial A(t)}{\partial k_3} = -A(0)\frac{K}{(K+1)^2} t * \exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.29}$$

$$\frac{\partial B(t)}{\partial k_1} = A(0)\left[\frac{k_2}{(k_1+k_2)^2} + \left(\frac{K}{K+1}\right)\frac{k_2}{(k_1+k_2)^2}(-k_3 t)\right]\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.30}$$

$$\frac{\partial B(t)}{\partial k_2} = A(0)\left[\frac{-k_1}{(k_1+k_2)^2} + \left(\frac{K}{K+1}\right)\frac{-k_1}{(k_1+k_2)^2}(-k_3 t)\right]\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.31}$$

$$\frac{\partial B(t)}{\partial k_3} = -A(0) * \left(\frac{K}{K+1}\right)^2 t\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.32}$$

$$\frac{\partial C(t)}{\partial k_1} = A(0) k_3 t \frac{k_2}{(k_1+k_2)^2}\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.33}$$

$$\frac{\partial C(t)}{\partial k_2} = A(0) k_3 t \frac{-k_1}{(k_1+k_2)^2}\exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.34}$$

$$\frac{\partial C(t)}{\partial k_3} = A(0)\frac{K}{K+1} t * \exp\left(-\frac{k_3 K t}{K+1}\right) \tag{A.35}$$

Appendix B

ACCELERATION AND SENSITIVITY ANALYSIS OF LATTICE KINETIC MONTE CARLO SIMULATIONS USING PARALLEL PROCESSING AND RATE CONSTANT RESCALING - SUPPLEMENT

B.1 Analytical solution of the Chemical Master Equation

For a continuous-time Markov chain (CTMC) with n_{states} states, let $p(t)$ be the $n_{\text{states}} \times 1$ column vector of probabilities of each state, such that $p_i(t)$ is the probability that the system is in state i at time t, and for all $t \geq 0$

$$\sum_{i=1}^{n_{\text{states}}} p_i(t) = 1. \tag{B.1}$$

The Chemical Master Equation gives the time evolution of the system according to the system of linear ordinary differential equations as

$$\frac{dp}{dt} = Q \cdot p. \tag{B.2}$$

The initial condition specifies $p(0)$. Q is the $n_{\text{states}} \times n_{\text{states}}$ matrix generator of the Markov chain and consists of all information relating to the event propensities. The time-dependent solution is

$$p(t) = \exp(Qt) \cdot p(0). \tag{B.3}$$

For a system with a finite number of states, the rank of Q is less than n_{states} due to the conservation constraint of Equation B.1. Therefore, the nullity of Q is at least 1 and there exists a non-trivial solution to the steady state equation

$$0 = Q \cdot p_{\text{ss}}. \tag{B.4}$$

p_{ss} is the vector of stationary probabilities for each of the states and is used to compute steady state properties. If there is only one recurrent class in Q, then the steady state

is unique. There will only be one recurrent class if any state can transition to any other state through a sequence of available reactions. For applications in chemical kinetics, this requires that the user specify a sufficient number of reaction types to access all available states from an arbitrary initial state. Steady state properties of Markov chains are discussed extensively in the mathematics literature[192].

Suppose there is a property f of each state of interest. Let f be a $1 \times n_{\text{states}}$ row vector for which f_i is the value of f in state i. The expected value of f at time t is given by

$$E\left[f\left(t\right)\right] = f \cdot p\left(t\right). \tag{B.5}$$

The reaction propensities, and therefore the generator Q may depend on parameters $\theta = [\theta_1, \theta_2, \dots]$, so that

$$Q = Q\left(\theta_1, \theta_2, \dots\right). \tag{B.6}$$

Time dependent sensitivities are computed by taking the derivative of Equation B.5, where all parameters other than θ_j are held fixed.

$$\frac{\partial E\left[f\left(t\right)\right]}{\partial \theta_j} = f \cdot \frac{\partial p\left(t\right)}{\partial \theta_j} + \frac{\partial f}{\partial \theta_j} \cdot p\left(t\right) \tag{B.7}$$

The second term on the right hand side is calculated as an expectation. The first term is expressed in terms of the generator by taking the derivative of Equation B.3 and differentiating the matrix exponential.

$$\frac{\partial p\left(t,\theta\right)}{\partial \theta_j} = \frac{\partial \exp\left(Qt\right)}{\partial \theta_j} \cdot p\left(0\right) \tag{B.8}$$

$$\frac{\partial p\left(t,\theta\right)}{\partial \theta_j} = \left[\int_0^1 \exp\left(\alpha Q\left(\theta\right)t\right) \frac{\partial Q\left(\theta\right)t}{\partial \theta_j} \exp\left(\left(1-\alpha\right)Q\left(\theta\right)t\right)\, d\alpha\right] p\left(0\right) \tag{B.9}$$

An analogous formula for the sensitivities of the steady state probabilities is derived in Hashemi et al.[82]

B.1.1 Autocorrelation and variance

Correlations between data at different time points ($t_1 \leq t_2$) are expressed in terms of the autocorrelation function.

$$\hat{c}\left(t_1, t_2\right) = \frac{\operatorname{cov}\left(f\left(t_1\right), f\left(t_2\right)\right)}{\sqrt{\operatorname{var}\left(f\left(t_1\right)\right) * \operatorname{var}\left(f\left(t_2\right)\right)}} \tag{B.10}$$

We derive an analytical expression for the autocorrelation function in terms of the generator. The numerator of Equation B.10 is expressed as

$$\operatorname{cov}\left(f\left(t_1\right), f\left(t_2\right)\right) = E\left[f\left(t_1\right) f\left(t_2\right)\right] - E\left[f\left(t_1\right)\right] E\left[f\left(t_2\right)\right]. \tag{B.11}$$

Equation B.5 allows us to evaluate $E\left[f\left(t\right)\right]$. We evaluate the quantity $E\left[f\left(t_1\right) f\left(t_2\right)\right]$ by enumerating the possible states i at t_1 and j at t_2. We use the variable $s\left(t\right)$ to give us the index of the state the trajectory is in at time t. Similarly, we use s_i to be a $n_{\text{states}} \times 1$ column vector of 0s except for a 1 at entry i. Our notation is such that $p_i\left(t\right) = s_i^T \cdot p\left(t\right)$.

$$E\left[f\left(t_1\right) f\left(t_2\right)\right] = \sum_{i=1}^{n_{\text{states}}} \sum_{j=1}^{n_{\text{states}}} f_j \left[p_j\left(t_2\right) | s\left(t_1\right) = i\right] f_i p_i\left(t_1\right) \tag{B.12}$$

The probability $p_j\left(t_2\right)$ is conditional on the system being in state i at time t_1.

$$= \sum_{i=1}^{n_{\text{states}}} \sum_{j=1}^{n_{\text{states}}} f_j \left[s_j^T \cdot \exp\left(Q\left(t_2 - t_1\right)\right) \cdot s_i\right] f_i \left[s_i^T \cdot \exp\left(Q t_1\right) \cdot p\left(0\right)\right] \tag{B.13}$$

$$= \sum_{j=1}^{n_{\text{states}}} f_j s_j^T \cdot \exp\left(Q\left(t_2 - t_1\right)\right) \cdot \sum_{i=1}^{n_{\text{states}}} f_i s_i \cdot s_i^T \cdot \exp\left(Q t_1\right) \cdot p\left(0\right) \tag{B.14}$$

We remove the summations over states using

$$\sum_{j=1}^{n_{\text{states}}} f_j s_j^T = f \tag{B.15}$$

and

$$\sum_{i=1}^{n_{\text{states}}} f_i s_i \cdot s_i^T = \operatorname{diag}\left(f\right) \tag{B.16}$$

$$E\left[f\left(t_1\right) f\left(t_2\right)\right] = f \cdot \exp\left(Q\left(t_2 - t_1\right)\right) \cdot \operatorname{diag}\left(f\right) \cdot \exp\left(Q t_1\right) \cdot p\left(0\right) \tag{B.17}$$

In the case of $t_1 = t_2$, we have

$$E\left[f(t)^2\right] = f \cdot \text{diag}\left(f\right) \cdot \exp\left(Qt\right) \cdot p\left(0\right) \tag{B.18}$$

When the system is at steady state, $\text{var}\left(f\left(t_1\right)\right) = \text{var}\left(f\left(t_2\right)\right) = \text{var}\left(f\right)$ and $E\left[f(t_1)\right] = E\left[f\left(t_2\right)\right] = f_{ss}$. Plugging Equation B.17 into Equation B.10 and applying the steady state condition gives

$$\hat{c}\left(t_1, t_2\right) = \hat{c}\left(\Delta t\right) = \frac{f \cdot \exp\left(Q\Delta t\right) \cdot \text{diag}\left(f\right) \cdot p_{ss} - f_{ss}^2}{\text{var}\left(f\right)} \tag{B.19}$$

We express the matrix exponential using eigenvector decomposition as

$$\exp\left(Q\Delta t\right) = \sum_{i=1}^{n_{\text{states}}} \exp\left(\lambda_i \Delta t\right) v_i \cdot z_i. \tag{B.20}$$

λ_i is the ith largest eigenvalue, while v_i and z_i are the corresponding right and left eigenvalues of Q respectively. Given the structure of the generator matrix, $\lambda_1 = 0$. Suitably normalized, $v_1 = p_{ss}$ and $z_1 = [1, 1, \ldots 1]$, allowing the first term of the series to be expressed as

$$f \cdot \exp\left(\lambda_1 \Delta t\right) v_1 \cdot z_1 \cdot \text{diag}\left(f\right) \cdot p_{ss} = f \cdot p_{\text{ss}} \cdot [1, 1, \ldots 1] \cdot \text{diag}\left(f\right) \cdot p_{ss} = f_{ss}^2, \tag{B.21}$$

thus canceling the f_{ss}^2 term in the numerator of Equation B.19.

The time scales of the system are defined by the opposite reciprocals of the nonzero eigenvalues of Q. That is, $\tau_i = -\frac{1}{\lambda_i}$. The exponential relaxation is defined as $\tau_{\text{exp}} = -\frac{1}{\lambda_2}$. If Δt is sufficiently larger than τ_3, then the exponential terms in the third term and beyond in Equation B.20 will be negligible, leaving only the second term. The time dependence occurs only in the exponential term, leaving the asymptotic behavior as

$$\hat{c}\left(t_1, t_2\right) = \hat{c}\left(\Delta t\right) \approx \exp\left(-\frac{\Delta t}{\tau_{\text{exp}}}\right). \tag{B.22}$$

Data at t_1 and t_2 are decorrelated if $\hat{c}\left(\Delta t\right) \ll 1$. For practical use, it is typical to use a decorrelation time of about $\Delta t \approx 3\tau_{\text{exp}}$ for which $\hat{c}\left(3\tau_{\text{exp}}\right) < 0.05$. For additional discussion of autocorrelation function approximations for Markov Chains, the reader is referred to other text[14].

B.1.2 Variance of time averages

Suppose that on a given time interval of length t_b, N evenly spaced data points are used to estimate the time-averaged mean of f. Without loss of generality, assume that $E\left[f\left(t\right)\right] = 0$. The data $f\left(t_1\right), \ f\left(t_2\right), \ldots f\left(t_N\right)$ are averaged, where $t_i = t_1 + t_b * \frac{i-1}{N-1}$. The time averaged mean estimate is

$$[f]_{t_1,t_2} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f\left(t_i\right) \tag{B.23}$$

$$\text{var}\left([f]_{t_1,t_2}\right) = E\left[\left(\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f\left(t_i\right)\right)^2\right] \tag{B.24}$$

$$= E\left[\lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N}\sum_{j=1}^{N} f\left(t_i\right) f\left(t_j\right)\right] \tag{B.25}$$

$$= \lim_{N\to\infty} \frac{1}{N}\left[\frac{1}{N} \sum_{i=1}^{N} E\left[f(t_i)^2\right] + \frac{2}{N} \sum_{i=1}^{N}\sum_{j<i}^{N} E\left[f\left(t_i\right) f\left(t_j\right)\right]\right] \tag{B.26}$$

Divide by $\text{var}\left(f\right) = E\left[f(t)^2\right]$

$$\frac{\text{var}\left([f]_{t_1,t_2}\right)}{\text{var}\left(f\right)} = \lim_{N\to\infty} \frac{1}{N}\left[1 + 2\sum_{i=1}^{N-1}\left(1 - \frac{i}{N}\right)\hat{c}\left(t_b\frac{i}{N}\right)\right] \tag{B.27}$$

We plug in Equation B.22 and take the limit $N \to \infty$ by converting to an integral to get

$$\approx \frac{2}{t_b}\int_0^{t_b}\left(1 - \frac{t}{t_b}\right)e^{-\frac{t}{t_b}}dt \tag{B.28}$$

The integral expression analytically evaluates to

$$\frac{\text{var}\left([f]_{t_1,t_2}\right)}{\text{var}\left(f\right)} \approx \frac{2\left(\tilde{t} + e^{-\tilde{t}} - 1\right)}{\tilde{t}^2} \tag{B.29}$$

where $\tilde{t} = \frac{t_b}{\tau_{\text{exp}}}$. Equation B.29 appears in Equation 3.6 and Figure 3.1 of the main text.

B.1.3 Autocorrelation of batch means

Here, we derive the autocorrelation of batch means of adjacent batches at steady state. Without loss of generality, we assume that $E\left[f\left(t\right)\right]=0$. The batch means are defined as

$$f_1=\frac{1}{t_b}\int_0^{t_b} f\left(t\right)dt \tag{B.30}$$

$$f_2=\frac{1}{t_b}\int_{t_b}^{2t_b} f\left(t\right)dt \tag{B.31}$$

The covariance of f_1 and f_2 is

$$E\left[f_1f_2\right]=\frac{1}{t_b^2}E\left[\left(\int_0^{t_b} f\left(t\right)dt\right)\left(\int_{t_b}^{2t_b} f\left(t\right)dt\right)\right] \tag{B.32}$$

$$=\frac{1}{t_b^2}\int_0^{t_b}\int_{t_b}^{2t_b} E\left[f\left(t_1\right)f\left(t_2\right)\right]dt_1dt_2 \tag{B.33}$$

Now we use the autocorrelation function according to Equation B.22 to replace the expectation in the integrand

$$\approx\frac{\mathrm{var}\left(f\right)}{t_b^2}\int_0^{t_b}\int_{t_b}^{2t_b}\exp\left(-\frac{\left|t_2-t_1\right|}{\tau}\right)dt_1dt_2 \tag{B.34}$$

$$\approx\frac{\mathrm{var}\left(f\right)}{t_b^2}\int_0^{t_b}\int_0^{t_b}\exp\left(-\frac{s_1+s_2}{\tau}\right)ds_1ds_2 \tag{B.35}$$

$$\approx\frac{\mathrm{var}\left(f\right)}{t_b^2}\left(\int_0^{t_b}\exp\left(-\frac{s_1}{\tau}\right)ds_1\right)^2 \tag{B.36}$$

$$\approx\frac{\mathrm{var}\left(f\right)}{t_b^2}\left(\int_0^{t_b}\exp\left(-\frac{s_1}{\tau}\right)ds_1\right)^2 \tag{B.37}$$

$$\frac{E\left[f_1f_2\right]}{\mathrm{var}\left(f\right)}\approx\left(\frac{1}{\tilde{t}}\left(1-\exp\left(-\tilde{t}\right)\right)\right)^2 \tag{B.38}$$

Finally, we divide by Equation B.29 to get

$$\frac{E\left[f_1f_2\right]}{\mathrm{var}\left(\left[f\right]_{t_b}\right)}\approx\frac{\left(1-e^{-\tilde{t}}\right)^2}{2\left(\tilde{t}+e^{-\tilde{t}}-1\right)} \tag{B.39}$$

Equation B.39 appears as the green dashed line in Figure 3.4 of the main text.

B.1.4 Variance of rate estimates

The time-averaged estimate the rate is computed as the mean of the batch means. The confidence interval of this estimate depends on the correlations between the batch means as follows.

$$\mathrm{var}\left(\mathrm{mean}\left(B_1 \cup B_2\right)\right) = \mathrm{var}\left(\frac{1}{n_{\mathrm{traj}}}\frac{1}{n_{\mathrm{bpt}}-1}\sum_{i=1}^{n_{\mathrm{traj}}}\sum_{j=2}^{n_{\mathrm{bpt}}} f_{i,j}\right) \tag{B.40}$$

Batches from different trajectories are uncorrelated, so we choose an arbitrary trajectory i and compute

$$\mathrm{var}\left(\mathrm{mean}\left(B_1 \cup B_2\right)\right) = \frac{1}{n_{\mathrm{traj}}}\mathrm{var}\left(\frac{1}{n_{\mathrm{bpt}}-1}\sum_{j=2}^{n_{\mathrm{bpt}}} f_{i,j}\right) \tag{B.41}$$

$$= \frac{1}{n_{\mathrm{traj}}}\frac{1}{\left(n_{\mathrm{bpt}}-1\right)^2}\left[\sum_{j=2}^{n_{\mathrm{bpt}}}\mathrm{var}\left(f_{i,j}\right) + 2\sum_{j_1=2}^{n_{\mathrm{bpt}}}\sum_{j_2=2}^{j_1-1}\mathrm{cov}\left(f_{i,j_1}, f_{i,\ j_2}\right)\right] \tag{B.42}$$

We use the definition that $\mathrm{cov}\left(f_{i,j_1}, f_{i,\ j_2}\right) = \mathrm{ACF}\left(f_{i,j_1}, f_{i,\ j_2}\right) * \mathrm{var}\left([f]_{\Delta t_{\mathrm{batch}}}\right)$ to simplify

$$= \frac{\mathrm{var}\left([f]_{\Delta t_{\mathrm{batch}}}\right)}{n_{\mathrm{traj}}\left(n_{\mathrm{bpt}}-1\right)}\left[1 + 2\sum_{j_1=2}^{n_{\mathrm{bpt}}}\sum_{j_2=2}^{j_1-1}\mathrm{ACF}\left(f_{1,j_1}, f_{1,\ j_2}\right)\right] \tag{B.43}$$

Assuming an exponential functional form for the autocorrelation, we substitute $\mathrm{ACF}\left(f_{1,j_1}, f_{1,\ j_2}\right) = \mathrm{ACF}(\Delta t_{\mathrm{batch}})^{|j_2 - j_1|}$, to obtain

$$= \frac{\mathrm{var}\left([f]_{\Delta t_{\mathrm{batch}}}\right)}{n_{\mathrm{batches}}}\left[1 + 2\sum_{\Delta j=1}^{n_{\mathrm{bpt}}-2}\left(1 - \frac{\Delta j}{n_{\mathrm{bpt}}-1}\right)\mathrm{ACF}^{\Delta j}\right] \tag{B.44}$$

The term outside the brackets of Equation B.44 is the naïve estimate of the variance of the mean. The term inside the brackets corrects for the effect of correlation between batches.

B.2 Analysis of A→B model

B.2.1 Analytical solution for A→B model

For a single site, there are n_{states}= 3 states: *, A*, and B*. The generator is

$$Q = \begin{bmatrix} -(k_1 P_A + k_{-3} P_B) & k_{-1} & k_3 \\ k_1 P_A & -(k_{-1} + k_2) & k_{-2} \\ k_{-3} P_B & k_2 & -(k_{-2} + k_3) \end{bmatrix} \tag{B.45}$$

Eigenvalue analysis yields a value of $\tau_{\text{exp}} = 0.4624\ s$. The initial condition is a bare lattice, so that

$$p(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{B.46}$$

The rate at each state is

$$f = \begin{bmatrix} -k_6 & 0 & k_5 \end{bmatrix} \tag{B.47}$$

Given these system specifications and the analytical formulas derived above, we obtain the analytical results shown in Figures 3.3-3.6. We also find that $f_{ss} = 0.1453$ and $\text{var}(f) = 0.0370$.

The numerical validation in Figure 3.1 in the main text was performed by running 1000 trajectories of length $t_f = 60\tau_{\text{exp}} = 27.744s$. The first $10\tau_{\text{exp}}$ of the trajectories was discarded so that only steady state data remained. For $n_{\text{traj}} = 1$, average rates were computed for each trajectory by time averaging the rate over the first $1.0\tau_{\text{exp}}$, $5.0\tau_{\text{exp}}$, $10.0\tau_{\text{exp}}$, and $50.0\tau_{\text{exp}}$ to create a data set for each value of $\tilde{t} = 1.0,\ 5.0,\ 10$, and 50. Data for $n_{\text{traj}} = 10$ was obtained by grouping the $n_{\text{traj}} = 1$ data in sets of 10 and taking the average. The variance of the data set for each combination of n_{traj} and $\tilde{t}$ was computed. Confidence intervals were computed with statistical bootstrapping.

Table B.1: Convergence test for the rescaling pare. r_{ss} is the estimated steady state rate for the A→B model. The CPU time accounts for all trajectories simulated to the same termination time of $t_f = 96.12\ s$.

$N_{\min}$	r_{ss} (site^{-1}s^{-1})	CPU time (s)
1	0.1337 ± 0.0002	532
10	0.1433 ± 0.0003	1020
50	0.1447 ± 0.0003	3289
100	0.1449 ± 0.0002	6562

Table B.2: Lattice size convergence calculations for the water-gas shift system. The quantity of each type of site in the lattice as well as the computed turnover frequency are shown.

Top	Bridge	Fcc hollow	Hcp hollow	Turnover Frequency (mol/site/sec)
200	600	200	200	76.0 ± 1.4
450	1350	450	450	77.9 ± 0.9
800	2400	800	800	78.9 ± 0.7

B.2.2 Determination of rate constant rescaling cutoff

Here we test the effect that different values of $N_{\min}$ have on the rate predictions and computational cost of the rate constant rescaling procedure. Results are shown in Table B.2. The error scales inversely with $N_{\min}$, while the computational cost scales linearly. We choose $N_{\min} = 50$ as a good tradeoff between accuracy and computational cost.

B.3 Additional data for the WGS System

To assess the effect of lattice size on the computed turnover frequencies for the water-gas shift system, simulations are performed with several different lattices of increasing size. The results, shown in Table B.2, indicate that the lattice size effect is small beyond 200 top sites, the lattice size used throughout this paper. They therefore should not affect the performance of the rate constant rescaling procedure or the physical interpretations made about the WGS system.

In Figure B.1, we plot the reaction frequencies at each iteration of the steady state convergence for the WGS reaction.

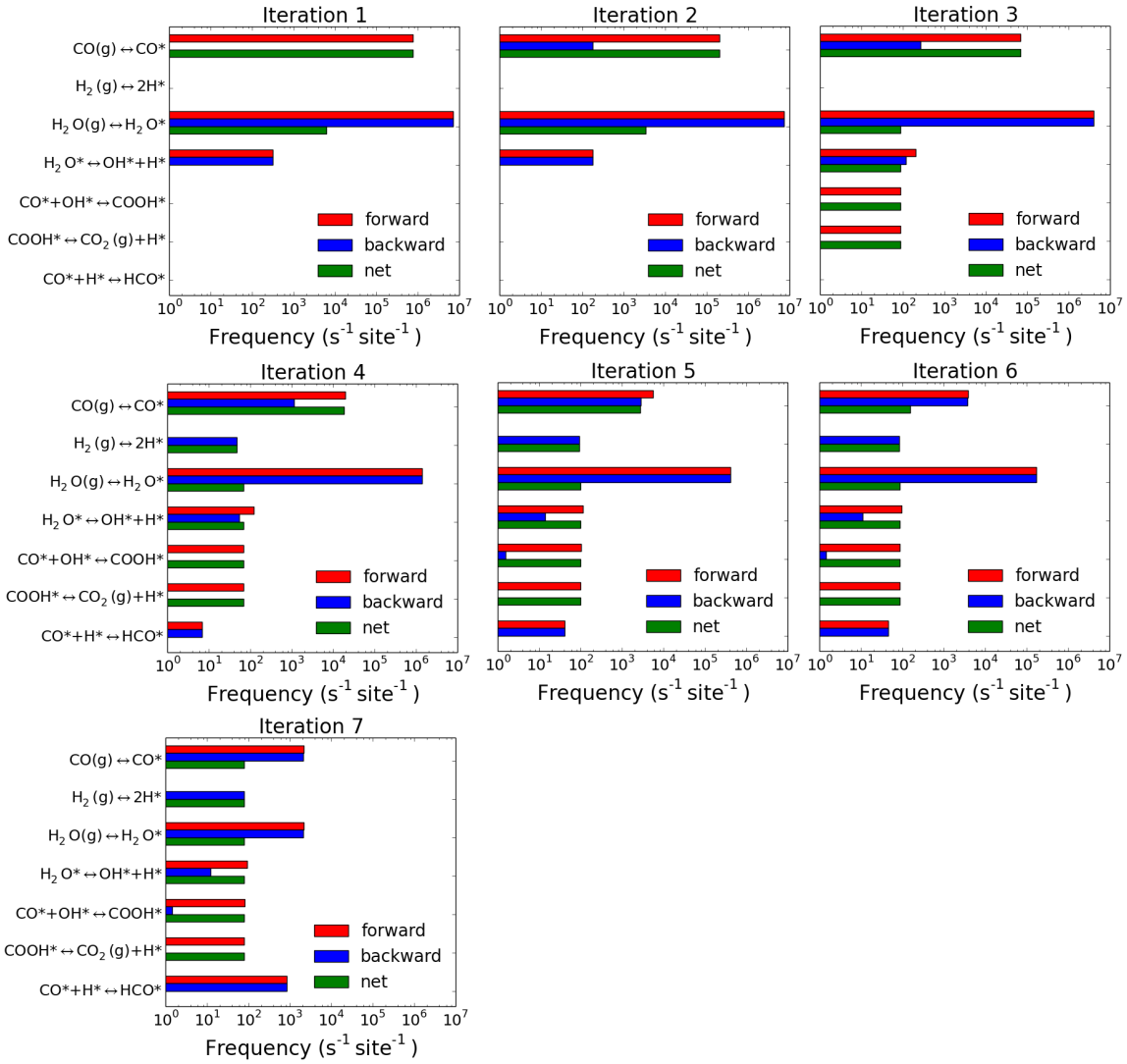


Figure B.1: Reaction frequency diagrams for each iteration of the WGS reaction simulation. As the iterations progress, H_2O adsorption and CO adsorption are scaled-down so that the reactions are less frequent. The simulation then samples on longer time scales so that the slower reactions are sampled.

Appendix C

OPTIMIZATION OF TRANSITION METAL CATALYST FACET STRUCTURE: APPLICATION TO THE OXYGEN REDUCTION REACTION - SUPPLEMENT

C.1 Reactions and Gibbs Energies vs. Coordination Number

The ORR reaction network consists of four elementary steps, Equations C.1-C.4.

$$O_2\,(g) + H^+ + e^- \rightarrow OOH^* \tag{C.1}$$

$$OOH^* + H^+ + e^- \rightarrow H_2O\,(l) + O^* \tag{C.2}$$

$$O^* + H^+ + e^- \rightarrow OH^* \tag{C.3}$$

$$OH^* + H^+ + e^- \rightarrow H_2O(l) \tag{C.4}$$

Two rate-determining steps have been identified: *OOH formation (Equation C.1) and *OH removal (Equation C.4)[152, 180]. *OH and *OOH are the key surface intermediates whose interactions with the catalyst determine its activity. Their binding energies have been computed on a wide range of metals and structures by Calle-Vallejo et al.[31, 28]

The number of evaluations of the objective function, which describes the catalyst activity for each structure, during optimization can be on the order of millions or more. The enormous computational burden of DFT precludes its use in optimization and therefore surrogate models are employed. Here we outline how the GCNs of a structure are used to achieve this. Using the data of Calle-Vallejo et al. for each metal, we propose the structural relationship describing the binding energies of *OH and *OOH vs. the GCN (Equations C.5 and C.6).

$$\Delta E_{OH}^{ads} = a_{OH} \cdot GCN + \beta_{OH} \tag{C.5}$$

$$\Delta E_{OOH}^{ads} = a_{OOH} \cdot GCN + \beta_{OOH} \tag{C.6}$$

Table C.1: Structural parameters for *OH and *OOH binding energies on Pt and Au, as used in Equations C.5 and C.6. All values are in units of eV.

Metal	a_{OH}	β_{OH}	a_{OOH}	β_{OOH}
Pt	0.19	-3.82	0.17	-2.34
Au	0.12	-3.09	0.12	-1.41

The estimated parameters for Pt and Au are shown in Table C.1. A large(small) value of α indicates high(low) degree of structure sensitivity while more (less) negative values of β indicate strong(weak) binding metals.

Knowledge of the *OH and *OOH binding energies allows us to approximate the overall activity from the Gibbs free energies of the two rate-determining steps. We use the standard approach taken in the electrochemistry literature, which uses the reversible hydrogen electrode[27, 31, 152].

The Gibbs free energies for all species, except O_2, are computed according to

$$G = E_{\mathrm{DFT}} + \mathrm{ZPE} - TS + E_{\mathrm{solv}} \tag{C.7}$$

and the Gibbs energy of O_2(g) is determined from the net thermodynamics according to

$$G_{O_2}(g) = 2 * \left(G_{\mathrm{H_2O(l)}} - G_{\mathrm{H_2}}\right) - \Delta G_{\mathrm{ORR}} \tag{C.8}$$

where $\Delta G_{\mathrm{ORR}} = -4.92$ eV. The terms on the right hand side of Equation C.7 are the DFT-calculated energy, the zero-point energy correction, the entropic contribution, and the solvation energy, respectively. Their values are taken from Calle-Vallejo et al. and are given in Table C.2. For a given cell voltage U, the Gibbs energies of reactions in Equations C.1 and C.4 are

$$\Delta G_1\left(U\right) = G_{*\mathrm{OOH}} - G_{\mathrm{O_2}}(g) + Ue \tag{C.9}$$

$$\Delta G_4\left(U\right) = G_{\mathrm{H_2O}}(l) - G_{*\mathrm{OH}} + Ue \tag{C.10}$$

Table C.2: Parameters used to compute species Gibbs free energies at a temperature of T= 298.15 K. All energies are in eV. The DFT energy of an adsorbed specie is the sum of the gas-phase energy and the binding energy. All values are taken from the supplementary information of Calle-Vallejo et al.[28] except for the gas-phase energies of OH and OOH, which are recalculated.

Species	E_{DFT}	ZPE	TS	E_{solv}
H_2(g)	-6.76	0.27	0.404	0
H_2O(l)	-14.22	0.574	0.583	-0.087
OH*	-7.53+ΔE_{OH}	0.332	0	-0.575
OOH*	-13.26+ΔE_{OOH}	0.480	0	-0.480

The second and third elementary steps are never rate limiting and are not considered further. The largest of the Gibbs free energies of reaction (Equations C.9 and C.10) is taken to be the Gibbs free energy of activation for the ORR process.

$$\Delta G_{\mathrm{ORR}} = \max\left(\Delta G_1(U), \Delta G_4(U)\right) \tag{C.11}$$

C.2 DFT Calculations

C.2.1 Validation of Surface Energy Model

In order to validate Equation 4.4, DFT calculations are performed. Formation energy is referenced to the bulk metal.

$$E_{\mathrm{form}} = E_{\mathrm{slab}} - n_{\mathrm{slab}} * E_{\mathrm{bulk}} \tag{C.12}$$

E_{slab} is the energy of the defected slab, n_{slab} indicates the number of atoms in the slab, and E_{bulk} is the energy per atom in the bulk.

DFT- and surrogate model-computed surface energies are compared in Figure C.1. The performance of the surrogate model is excellent when the value of E_{coh} is adjusted to fit the DFT data. Deviation from the experimental value is expected given the inherent errors in DFT and the approximate nature of Equation 4.4.

DFT calculations are implemented in the VASP software[112, 114, 113, 115]. Each surface is modeled as a six-layer slab with p(4x4) periodicity in the horizontal dimensions and a 15 Å vacuum slab in the z dimension. The sixth layer may be

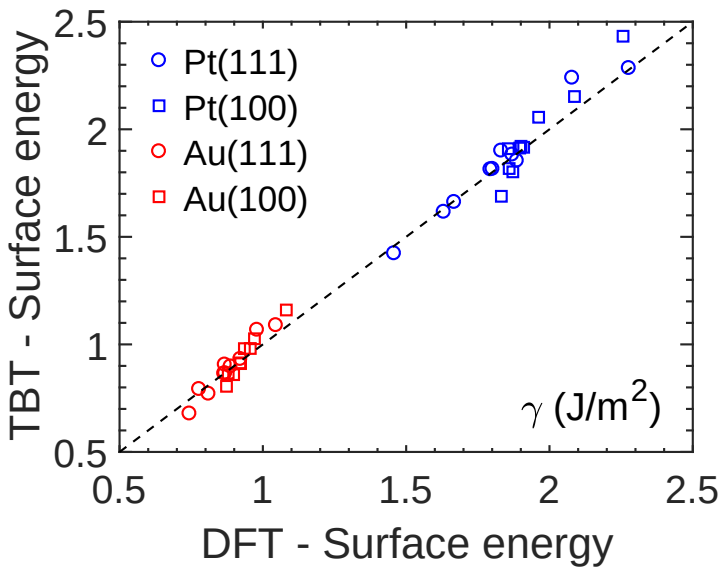


Figure C.1: Parity plot of surface energies (γ) computed with the tight-binding model (TBT) (Equations 4.3 and 4.4) vs. DFT. The data set includes structures of defected (111) (circles) and (100) (square) planes of Pt (blue) and Au (red).

incomplete to describe defects. The bottom two layers are frozen while the top layers relax. The Perdew, Burke, and Ernzerhod (PBE)[161] exchange correlation functional, PAW pseudopotentials[18, 116], and a Monkhorst-Pack k-point mesh of 3x3x1 are used. Ground state energies are calculated using a plane wave basis set with a kinetic energy cutoff of 400 eV, a smearing width of 0.1 eV, an electronic energy threshold of 10^{-4} eV, and a force tolerance of 0.04 eV/Å. Lattice constants of 3.97 Å and 4.16 Å are used for Pt and Au, respectively, and are computed by relaxing the size of the fcc unit cell. The formation energy of the bottom part of the slab, modeling the bulk, is taken into account by subtracting half of the formation energy of the ideal 6-layer slab.

The experimental, DFT-computed, and fitted values of the cohesive energies are shown in Table C.3. DFT values are computed by taking the difference between the bulk energy and that of an isolated metal atom in the gas phase. DFT under predicts the cohesive energies of Pt and Au, consistent with the findings of Janthon et al.[95] The values of the cohesive energies resulting from a fit of Equation 4.4 to the DFT data are even lower. Analysis of how DFT predicts surface energies rationalizes

Table C.3: Comparison of cohesive energies obtained by experiments, DFT, and fits of Equation 4.4 to DFT data. Experimental values are taken from Kittel[109].

Cohesive Energy, E_{coh} (eV)	Pt	Au
Experimental[109]	5.84	3.81
DFT	5.56	3.04
Equation 4.4 fit	4.52	2.36

Table C.4: Comparison of surface energies from experiments and DFT. Experimental values are taken from Tyson et al.[218] Systematic under prediction by DFT can be observed, consistent with Janthon et al.[95]

Surface energy, γ (J/m^2)	Pt(111)	Au(111)
Experimental[218]	2.489	1.506
DFT	1.458	0.744

these low fitted values. Table C.4 compares the experimental and DFT values of the surface energies of Pt(111) and Au(111). DFT under predicts surface energies even more severely than it under predicts cohesive energies. Janthon et al. tabulates similar data for a variety of functionals and reached the same conclusion. Due to this under prediction of surface energy, low fitted values of the cohesive energies relative to the experimental values are needed to make Equation 4.4 consistent with the DFT values.

In order to assess the range of interactions between different sites, we consider a simple case and compute the vacancy formation energies of point defects (i.e. a single-atom vacancy) at varying distances in a 5-layer $p(5 \times 5)$ slab of Pt(111). For each pair of defects, we compute the excess energy, which we define as the difference in the formation energy of both defects and twice the vacancy formation energy for a single point defect. The results of these calculations are shown in Figure C.2. We find that the interactions of vacancies on each other are negligible beyond first nearest neighbor. The short range of this interaction justifies the use of coordination number for formation energies, as well as the lack of coupling between sites in our model for activity. For large defects, the active site is typically beyond first nearest neighbor and as a result, we do not expect significant interaction. Coupling between sites could be included by performing DFT calculations on a set of training structures with multiple

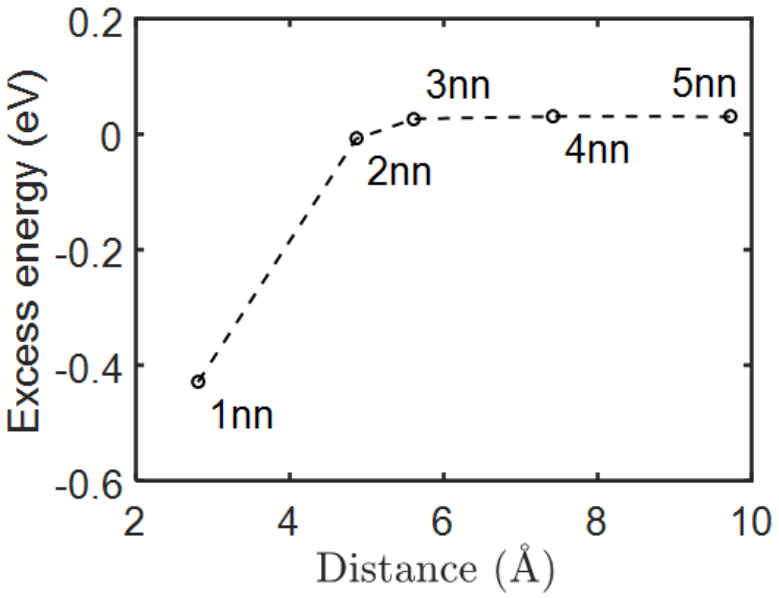


Figure C.2: Convergence of excess vacancy formation energies for pairs of defects. Excess energy on the y-axis is defined as the vacancy formation energy of a pair of defects, minus twice the vacancy formation energy of a single point defect. The x-axis is the distance between the defects, i.e., the atoms removed to create the vacancies. Points are labeled as 1^{st} through 5^{th} nearest neighbors in the fcc (111) lattice.

defects, and regressing a cluster expansion Hamiltonian, as has been done for oxygen adlayers on metal surfaces[190, 234, 239].

C.2.2 Validation of Active Sites

In order to assess the accuracy of the GCN model, we have computed the *OH and *OOH binding energies of the predicted defected sites using the same VASP parameters as above. This approach provides an inexpensive method to refine the predictions of the screening model, following recent advances in a hierarchical model development[210, 220] (in brief, a screening model is used first, important parameters or defects in our case are identified, and higher level theory is employed to refine the results). The structures for DFT are shown in Figure C.3. The metal slabs consist of four complete layers with additional adatoms for the defect layer. Site (e) is modeled in DFT as a 12-layer (211) surface with an additional atom on the step (see Figure

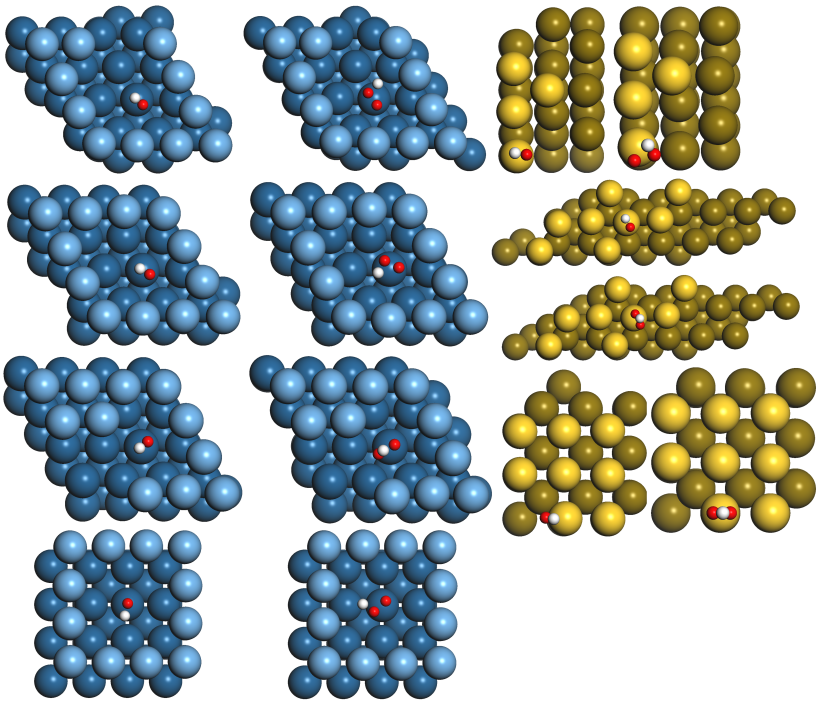


Figure C.3: Active site structures used for DFT validation. Slabs with adsorbed *OH are shown on the left, while slabs with *OOH adsorbed are shown on the right. The rows show active sites a-g, ordered top to bottom.

C.3). The binding energies are tabulated in Table C.5, while the predicted activities are shown in Figure C.4.

DFT computed binding energies differ from those predicted using the generalized coordination number by between 0.05 eV and 0.4 eV. Many are less than 0.1 eV and most are within 0.3 eV, i.e., nearly within DFT error. The errors are only significant for two out of the seven active sites. For sites (a) through (e), the activities predicted using DFT versus the GCN still offer significant improvement relative to the ideal surface. Site (c), for example, offers an improvement relative to a Pt(111) top site of about 30. In contrast, the activities of sites (a) and (b) are actually underpredicted by the GCN volcano. Given the disordered arrangement of sites in our Pt(111) defected structure, errors for different sites likely compensate for each other. For sites (f) and (g), DFT predicts the activity to be less than the volcano plots would predict due to weaker binding of *OOH.

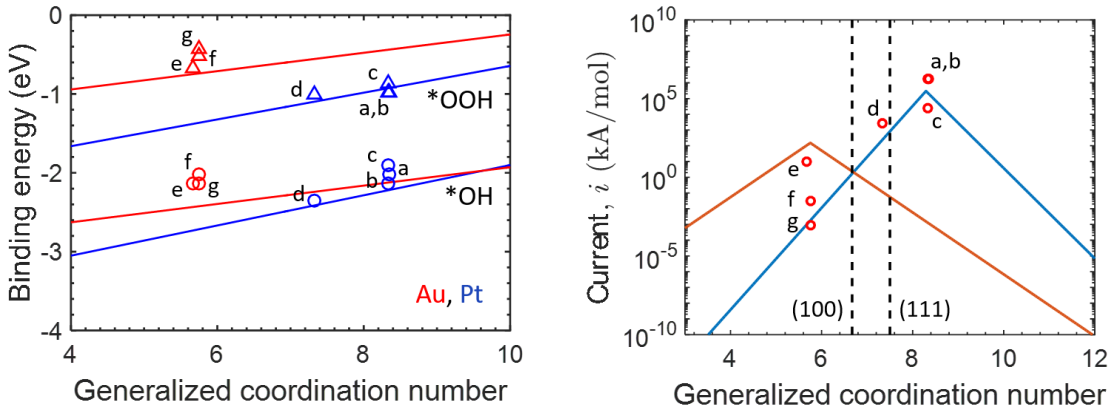


Figure C.4: (left) Graphical depiction of the data in Table C.5. Points indicate DFT computed *OH and *OOH binding energies for each active site. Triangles and circles denote *OOH and *OH binding energies, respectively. Lines indicate predictions based on the GCN, using Equations C.5 and C.6. (right) Points indicate activities predicted for each active site, using Equation 4.1 and DFT energies. Lines indicate activity predictions based on the GCN, which use Equations C.5-C.11, as well as Equation 4.1. Sites a-e are approximately within an order of magnitude of their predicted activity. Sites (f) and (g) are predicted to be less active due to weaker binding of *OOH.

Table C.5: Validation of binding energy predictions for the active sites encountered in optimization. $\Delta E^{\text{ads}}_{\text{OH, DFT}}$ is the binding energy of *OH (or *OOH) computed with DFT. $\Delta E^{\text{ads}}_{\text{OH, GCN}}$ is the binding energy predicted using Equation C.5 or Equation C.6. All energies are in eV. Errors are expressed as $\Delta\Delta E^{\text{ads}}_{\text{OH}} = \Delta E^{\text{ads}}_{\text{OH,DFT}} - \Delta E^{\text{ads}}_{\text{OH,GCN}}$ and similarly for *OOH.

Active Site	Metal	GCN	$\Delta E^{\text{ads}}_{\text{OH, DFT}}$	$\Delta\Delta E^{\text{ads}}_{\text{OH}}$	$\Delta E^{\text{ads}}_{\text{OOH,DFT}}$	$\Delta\Delta E^{\text{ads}}_{\text{OOH}}$
a	Pt	8.35	-2.017	0.200	-0.982	-0.058
b	Pt	8.33	-2.134	0.087	-0.979	-0.052
c	Pt	8.33	-1.907	0.315	-0.870	0.058
d	Pt	7.33	-2.349	0.064	-1.013	0.084
e	Au	5.67	-2.138	0.295	-0.672	0.077
f	Au	5.75	-2.019	0.405	-0.520	0.219
g	Au	5.75	-2.131	0.293	-0.431	0.309

C.3 Computational Acceleration

C.3.1 Structure modeling

A slab with dimensions $p(d_1 \times d_2)$ has $n_{\text{pos}} = 4 \times d_1 \times d_2$ lattice positions. A $n_{\text{pos}} \times n_{\text{pos}}$ adjacency matrix A is used, where $A_{i,j} = 1$ if the positions i and j are nearest neighbors and 0 otherwise. Each row of A has at most $\text{CN}_{\text{max}} = 12$ nonzero elements and therefore A is sparse, a trait that is exploited to accelerate computation. The column vectors of coordination numbers $\vec{\text{CN}}$ and generalized coordination numbers $\vec{\text{GCN}}$ for each atom are computed according to

$$\vec{\text{CN}} = A \cdot \vec{c} \tag{C.13}$$

$$\vec{\text{GCN}} = \frac{1}{\text{CN}_{\text{max}}} A \cdot \text{diag}(\vec{c}) \cdot \vec{\text{CN}} \tag{C.14}$$

$\text{diag}(\vec{c})$ indicates a diagonal matrix whose diagonal elements are the elements of $\vec{c}$. The coordination numbers and GCNs of the atoms in the bottommost layer are 12 to account for the bulk metal.

C.3.2 Choosing adjacent sites for energy minimization

We efficiently list all available pairs of occupancies and adjacent vacancies as follows. The matrices A_{var} and $\vec{c}_{\text{var}}$ are taken as A and $\vec{c}$ with only the rows and columns corresponding to the atoms in the top layer. Then, the matrix A' is computed as

$$A' = \text{diag}(\vec{c}_{\text{var}}) \cdot A_{\text{var}} \cdot (I_{n_{\text{var}}} - \text{diag}(\vec{c}_{\text{var}})) \tag{C.15}$$

so that $A'_{i,j} = 1$ if i is an occupied site and j is an adjacent unoccupied site. Otherwise, $A'_{i,j} = 0$. $I_{n_{\text{var}}}$ denotes the $n_{\text{var}} \times n_{\text{var}}$ identity matrix and n_{var} is the number of atom positions in the top layer. A random non-zero term of A' is chosen so that the candidate event is the transfer of an atom from site i to site j.

C.3.3 Normalizing activity and stability

We non-dimensionalize surface energy and activity to achieve better numerical behavior. Energy differences are scaled by the energy required to break a metal-metal bond, while activity is normalized by the current density of an active site.

$$\mathcal{E} = \frac{E_{\text{form}}}{E_{\text{coh}}/\text{CN}_{\text{max}}} \tag{C.16}$$

$$\mathcal{I} = -\frac{\sum_k i_k}{s * i_{\text{max}}} \tag{C.17}$$

The negative sign in Equation C.17 is employed because activity must be maximized rather than minimized. i_{max} is the value of i at the peak of the volcano curve for Pt or Au. s is a scaling factor used to improve numerics ($s = 0.5$ is used here).

The energy of all the bonds a metal atom has in the bulk is the cohesive energy of the solid, E_{coh}. Given that it has CN_{max} bonds in this case, the quantity $E_{\text{coh}}/\text{CN}_{\text{max}}$ is the average bond strength of metal-metal bonds. Metal-metal bonds are broken and formed during optimization, so the average bond strength is an appropriate normalization factor.

Similarly, for current density, an appropriate scaling factor approximates changes during optimization. The movement of atoms may destroy or create an active site, in which case the current density of an active site is desired. We use the value of current density at the volcano peak, i_{max} to approximate the activity of an active site. However, the active site may not be quite as active as the volcano peak, and changes in structure may not completely destroy or create an active site. Therefore, the factor $s = 0.5$ scales down the normalization factor to changes in activity that would be more common. Additionally, changes in dimensionless surface energy and current density must be comparable, so that they can be subjected to the same cooling schedule and be linearly mixed.

C.4 Metropolis Algorithm

C.4.1 Algorithmic details

The multi-objective optimization uses a Metropolis approach with simulated annealing. A candidate move would change the objective function by an amount $\Delta\mathcal{F}$. If $\Delta\mathcal{F} < 0$, the move is accepted; if $\Delta\mathcal{F} > 0$, the move is accepted if

$$\exp\left(-\frac{\Delta\mathcal{F}}{T_{\text{metro}}}\right) > \text{RN}\left([0,1]\right) \tag{C.18}$$

where $\text{RN}\left([0,1]\right)$ is a random number taken from the uniform probability distribution between 0 and 1 and T_{metro} is the dimensionless Metropolis temperature. T_{metro} is given by a linear cooling schedule[153]

$$T_{\text{metro}}\left(\text{step}\right) = T_{\text{max}}\left(1 - \frac{\text{step}}{\text{step}_{\text{max}}}\right) \tag{C.19}$$

$T_{\text{max}} = 0.7$ is used. step_{max} is the maximum number of optimization steps; a value of 450,000 is used.

During the energy minimization, a move consists of a top-layer atom in an occupied site moving to an adjacent unoccupied site. The move is accepted if and only if $\Delta\mathcal{E} < 0$ in a downhill energy-minimization approach. 45,000 steps are used. An example of optimization, demonstrating how activity and surface energy change as the simulation progresses through the multiobjective and subsequent energy-minimization, is shown in Figure C.5.

C.4.2 Optimization pseudocode

Initialization

1. Read parameters from an input file.
2. Build catalyst structure, e.g., Pt(111), Pt(100), Au(111), or Au(100).
3. Build adjacency matrix A of nearest neighbors.

Multi-objective Optimization (maximize $\mathcal{I}$ and minimize $\mathcal{E}$ simultaneously)

4. Initialize the catalyst slab with random occupancies in the top layer.

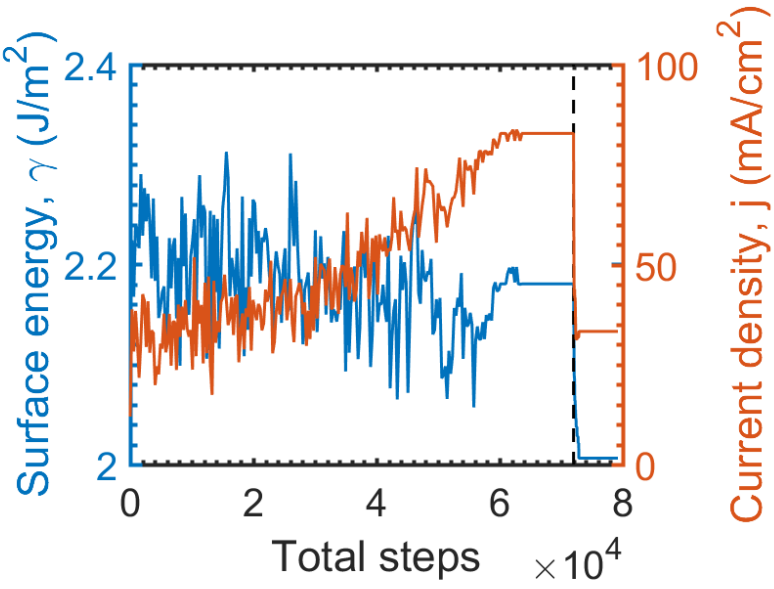


Figure C.5: Example of an optimization of a Pt(111) surface using an activity-weight of $\omega = 0.7$. Current density (Equation 4.2) increases while surface energy (Equation 4.3) decreases during the multiobjective optimization (left of the vertical dashed line). During the subsequent energy-only minimization (right of the vertical dashed line), the surface energy decreases, typically along with a decrease in activity (no uphill moves are accepted in this second level resulting in lack of noise).

5. Set $0 < \omega \leq 1$
6. Compute CNs and GCNs of all atoms using Equations C.13 and C.14 .
7. Compute $\mathcal{I}$ and $\mathcal{E}$ for the initial structure using Equations C.16 and C.17.
8. Compute $\mathcal{F}$ using Equation 4.5 .
9. Choose a random lattice position in the top layer and attempt to switch its occupancy.
10. Perform steps 6-8 for the proposed structure.
11. Choose temperature T_{metro} according to Equation C.19.
12. If Equation C.18 is satisfied, the new structure is accepted.
13. Repeat steps 9-12 for a step_{max} steps.
14. Output data for the structure encountered with the smallest value of $\mathcal{F}$.

Energy-only minimization (locally minimize $\mathcal{E}$)

15. Initialize structure from the multi-objective optimization.
16. Choose a random pair of adjacent occupied and unoccupied sites in the top layer using Equation C.15 and attempt to swamp them.
17. Compute $\mathcal{E}$ for the new structure.
18. If $\Delta\mathcal{E} < 0$, accept the new structure.
19. Repeat steps 16-18 for a step_{max} steps.

C.5 Pareto Plots

For a given metal-facet, a number of simulations are run to achieve a variety of optima. 250 simulations are found to be sufficient. Each simulation has a different random initial structure, random seed, and ω value. The k^{th} simulation uses a value of $\omega_k = \frac{k}{250}$, where $k = 1,\ 2, \ldots,\ 250$. The collection of 250 optima from all simulations forms a population of structures for which we compute the surface energies and current densities. The inherent randomness of the optimization results in some structures having superior properties to others. We identify the structures of interest to be those

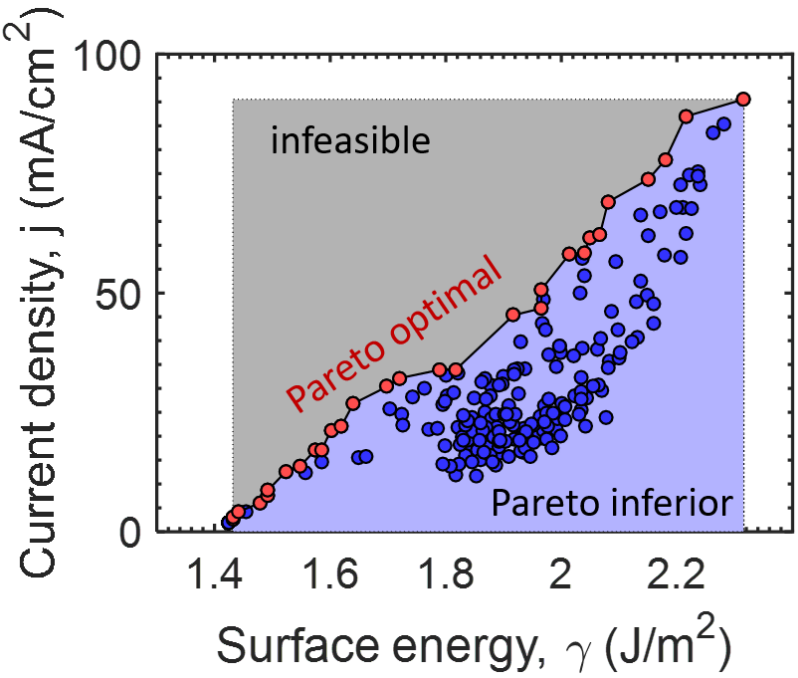


Figure C.6: The concept of Pareto optimality demonstrated using Pt(111) data. Data points correspond to structures resulting from the multiobjective optimization followed by subsequent energy minimization for various weights (ω). ω near 1 maximizes primarily activity and ω near 0 chiefly minimizes energy. The red points connected with the solid black line are the Pareto optimal points. A point p_1 is Pareto optimal if there is no other point p_2 for which $\gamma(p_2) < \gamma(p_1)$ and $j(p_2) > j(p_1)$.

with high current density and low surface energy. Figure C.6 demonstrates how the Pareto front is computed from a large number of calculations using different values of the activity weight ω.

C.6 Activity Fluctuations

Figure C.8 shows results for the other surfaces, analogous to Figure 4.3 in the main text. Figure C.5 shows the fluctuations of activity for a Pt(111) surface at a constant loading of 49% vacancies.

C.7 Suboptimal defects

Figure C.9 shows a Pt(111) surface with suboptimal defects. It has a surface energy of 1.80 J/m^2 and a current density of 0.048 mA/cm^2, and is therefore about

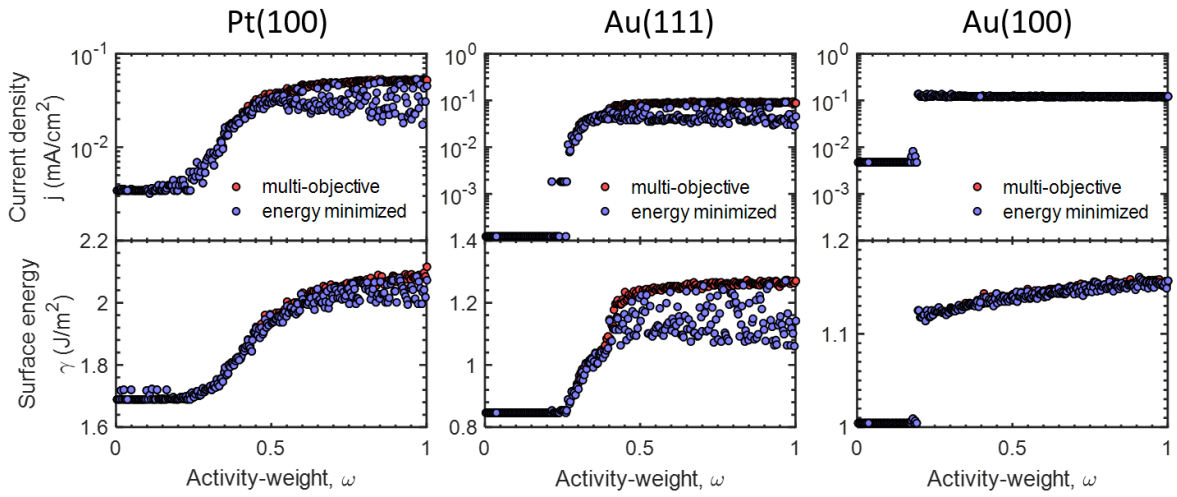


Figure C.7: Results of the optimization parameter sweep for Pt(100), Au(111), Au(100). Current densities (Equation 4.2) and surface energies (Equation 4.3) are plotted against activity-weight ω. Red points show values resulting from multiobjective optimization only. Blue points show data after the energy-only minimization step is also applied.

two orders of magnitude less active than Pt(111). This demonstrates conceptually that surface roughening can reduce the activity of a pristine surface and deliberate design of active sites is necessary to increase activity.

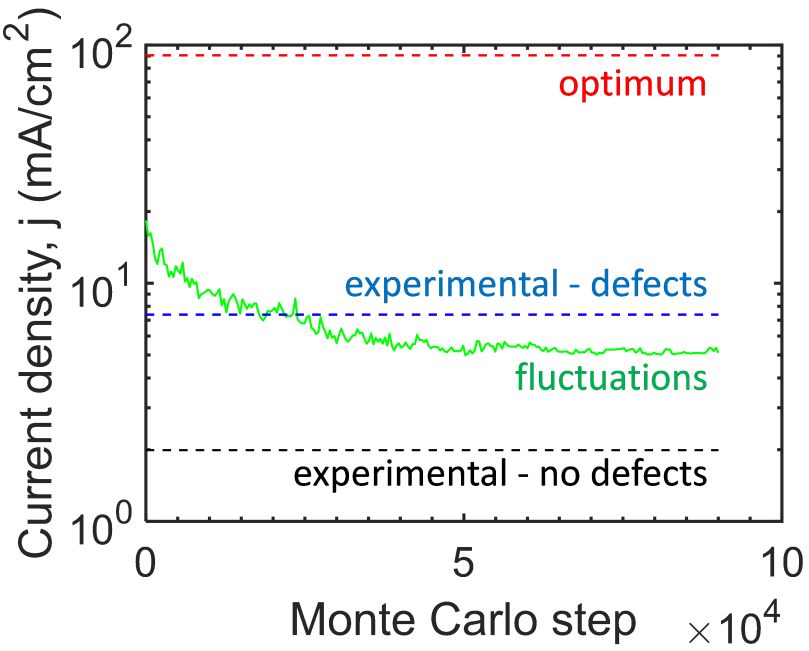


Figure C.8: Comparison of experimental activities with those computed through optimization. The red line marks the activity of the most active defected Pt(111) structure achieved from optimization. The blue and black lines show experimental data from Calle-Vallejo et al.[28] for Pt(111) with and without defects, respectively. The fluctuating green curve shows a trajectory for a Monte Carlo simulation in which a surface is initialized with a random half-monolayer of vacancies and allowed to reconstruct at finite temperature, which is gradually decreased. Surface energy is used as the objective function in a Metropolis Monte Carlo simulation.

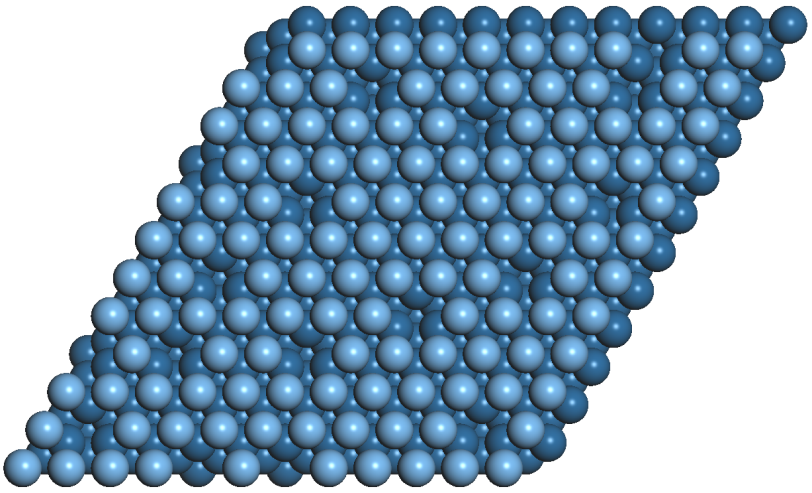


Figure C.9: Example of a Pt(111) surface in which the presence of defects decreases activity. Vacancies are present, but do not expose active sites in the second layer. The vacancies also decrease the GCNs of active sites on the top layer, decreasing their activity. As a result, the current density is about two orders of magnitude less than that of Pt(111).

Appendix D

AN ONLINE MACHINE LEARNING APPROACH TO OPTIMIZING ACTIVE SITE COUPLING - SUPPLEMENT

D.1 Site indexing and symmetry operators

We show mathematically how the rotational operator ($R_\theta(\sigma)$) permutes the site indices. Let the fractional coordinates of a Ni site be

$$f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}. \tag{D.1}$$

The Cartesian coordinates of the fcc(111) unit cell is

$$c = d_{\mathrm{Pt}} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \tag{D.2}$$

where d_{Pt} is the nearest neighbor distance between Ni sites, as imposed by the lattice constant of the Pt substrate. The rotation matrix is

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{D.3}$$

The fractional coordinates will then be transformed as

$$f^{\mathrm{new}} = c^{-1} \cdot R \cdot c \cdot f. \tag{D.4}$$

In our case, it becomes

$$c^{-1} \cdot R\left(\frac{2\pi}{3}\right) \cdot c = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \tag{D.5}$$

$$c^{-1} \cdot R\left(\frac{4\pi}{3}\right) \cdot c = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}. \tag{D.6}$$

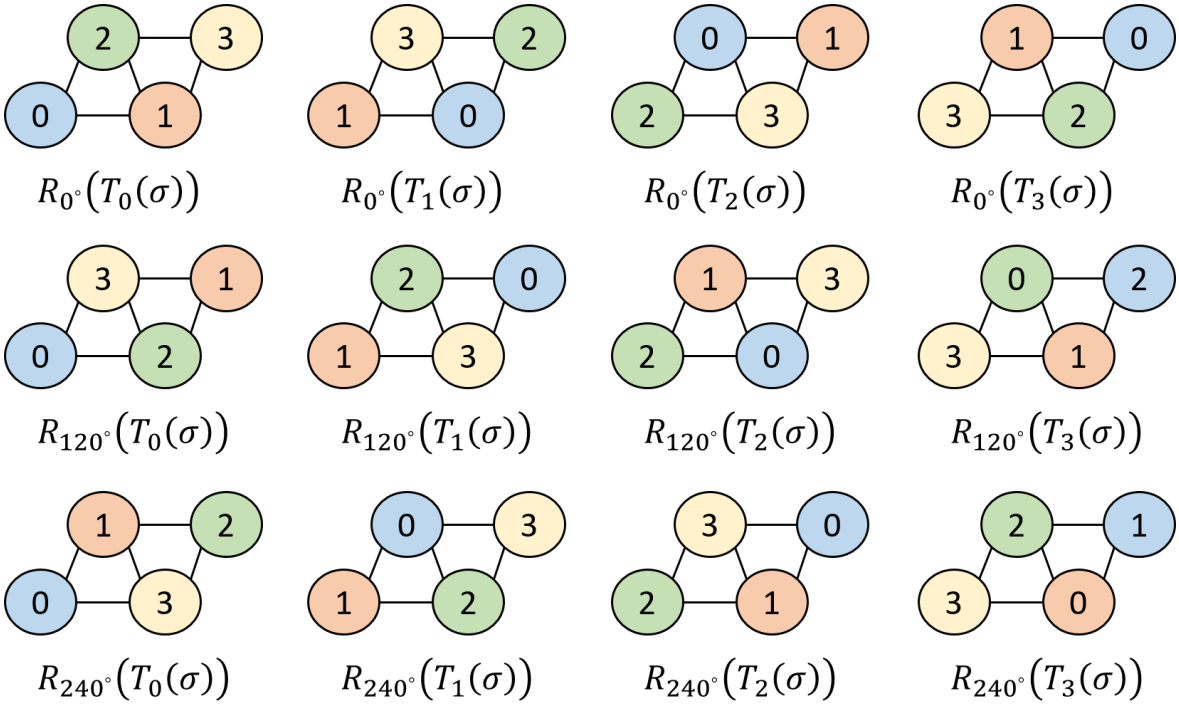


Figure D.1: All 12 translational and rotational symmetries of a $p(2\times2)$ fcc(111) lattice. Numbers indicate the indices of the Ni sites of the original lattice.

The new Ni site indices after $R\,(240°)$ are

$$f_1^{\text{new}} = (-f_1 - f_2) \mod d \tag{D.7}$$

$$f_2^{\text{new}} = f_1 \mod d \tag{D.8}$$

$$i = f_2 d + f_1 + 1. \tag{D.9}$$

An example of how the rotational and translational operators permute the site indices is shown in Figure D.1.

Appendix E

PERMISSIONS

AIP PUBLISHING LLC LICENSE TERMS AND CONDITIONS

Feb 01, 2018

This Agreement between Mr. Marcel Nunez ("You") and AIP Publishing LLC ("AIP Publishing LLC") consists of your license details and the terms and conditions provided by AIP Publishing LLC and Copyright Clearance Center.

License Number	4280541075276
License date	Feb 01, 2018
Licensed Content Publisher	AIP Publishing LLC
Licensed Content Publication	Journal of Chemical Physics
Licensed Content Title	Steady state likelihood ratio sensitivity analysis for stiff kinetic Monte Carlo simulations
Licensed Content Author	M. Núñez, D. G. Vlachos
Licensed Content Date	Jan 28, 2015
Licensed Content Volume	142
Licensed Content Issue	4
Type of Use	Thesis/Dissertation
Requestor type	Author (original article)
Format	Electronic
Portion	Excerpt (> 800 words)
Will you be translating?	No
Title of your thesis / dissertation	UNCERTAINTY QUANTIFICATION IN MULTISCALE STOCHASTIC MODELS OF CATALYTIC REACTIONS
Expected completion date	Feb 2018
Estimated size (number of pages)	150
Requestor Location	Mr. Marcel Nunez United States Attn: Mr. Marcel Nunez
Billing Type	Invoice
Billing Address	Mr. Marcel Nunez United States Attn: Mr. Marcel Nunez
Total	**0.00 USD**

Terms and Conditions

AIP Publishing LLC -- Terms and Conditions: Permissions Uses

AIP Publishing hereby grants to you the non-exclusive right and license to use and/or distribute the Material according to the use specified in your order, on a one-time basis, for the specified term, with a maximum distribution equal to the number that you have ordered. Any links or other content accompanying the Material are not the subject of this license.

1. You agree to include the following copyright and permission notice with the reproduction of the Material:"Reprinted from [FULL CITATION], with the permission of AIP Publishing." For an article, the credit line and permission notice must be printed

on the first page of the article or book chapter. For photographs, covers, or tables, the notice may appear with the Material, in a footnote, or in the reference list.

2. If you have licensed reuse of a figure, photograph, cover, or table, it is your responsibility to ensure that the material is original to AIP Publishing and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover, or table does not indicate that it was reprinted by AIP Publishing, with permission, from another source. Under no circumstances does AIP Publishing purport or intend to grant permission to reuse material to which it does not hold appropriate rights.
   You may not alter or modify the Material in any manner. You may translate the Material into another language only if you have licensed translation rights. You may not use the Material for promotional purposes.
3. The foregoing license shall not take effect unless and until AIP Publishing or its agent, Copyright Clearance Center, receives the Payment in accordance with Copyright Clearance Center Billing and Payment Terms and Conditions, which are incorporated herein by reference.
4. AIP Publishing or Copyright Clearance Center may, within two business days of granting this license, revoke the license for any reason whatsoever, with a full refund payable to you. Should you violate the terms of this license at any time, AIP Publishing, or Copyright Clearance Center may revoke the license with no refund to you. Notice of such revocation will be made using the contact information provided by you. Failure to receive such notice will not nullify the revocation.
5. AIP Publishing makes no representations or warranties with respect to the Material. You agree to indemnify and hold harmless AIP Publishing, and their officers, directors, employees or agents from and against any and all claims arising out of your use of the Material other than as specifically authorized herein.
6. The permission granted herein is personal to you and is not transferable or assignable without the prior written permission of AIP Publishing. This license may not be amended except in a writing signed by the party to be charged.
7. If purchase orders, acknowledgments or check endorsements are issued on any forms containing terms and conditions which are inconsistent with these provisions, such inconsistent terms and conditions shall be of no force and effect. This document, including the CCC Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

This Agreement shall be governed by and construed in accordance with the laws of the State of New York. Both parties hereby submit to the jurisdiction of the courts of New York County for purposes of resolving any disputes that may arise hereunder.

V1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

AIP PUBLISHING LLC LICENSE TERMS AND CONDITIONS

Feb 01, 2018

This Agreement between Mr. Marcel Nunez ("You") and AIP Publishing LLC ("AIP Publishing LLC") consists of your license details and the terms and conditions provided by AIP Publishing LLC and Copyright Clearance Center.

License Number	4280541477241
License date	Feb 01, 2018
Licensed Content Publisher	AIP Publishing LLC
Licensed Content Publication	Journal of Chemical Physics
Licensed Content Title	Stochastic averaging and sensitivity analysis for two scale reaction networks
Licensed Content Author	Araz Hashemi, Marcel Núñez, Petr Plecháč, et al
Licensed Content Date	Feb 21, 2016
Licensed Content Volume	144
Licensed Content Issue	7
Type of Use	Thesis/Dissertation
Requestor type	Author (original article)
Format	Electronic
Portion	Excerpt (> 800 words)
Will you be translating?	No
Title of your thesis / dissertation	UNCERTAINTY QUANTIFICATION IN MULTISCALE STOCHASTIC MODELS OF CATALYTIC REACTIONS
Expected completion date	Feb 2018
Estimated size (number of pages)	150
Requestor Location	Mr. Marcel Nunez United States Attn: Mr. Marcel Nunez
Billing Type	Invoice
Billing Address	Mr. Marcel Nunez United States Attn: Mr. Marcel Nunez
Total	0.00 USD

Terms and Conditions

AIP Publishing LLC -- Terms and Conditions: Permissions Uses

AIP Publishing hereby grants to you the non-exclusive right and license to use and/or distribute

the Material according to the use specified in your order, on a one-time basis, for the specified term, with a maximum distribution equal to the number that you have ordered. Any links or other content accompanying the Material are not the subject of this license.

1. You agree to include the following copyright and permission notice with the reproduction of the Material:"Reprinted from [FULL CITATION], with the permission of AIP Publishing." For an article, the credit line and permission notice must be printed on the first page of the article or book chapter. For photographs, covers, or tables, the notice may appear with the Material, in a footnote, or in the reference list.
2. If you have licensed reuse of a figure, photograph, cover, or table, it is your responsibility to ensure that the material is original to AIP Publishing and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover, or table does not indicate that it was reprinted by AIP Publishing, with permission, from another source. Under no circumstances does AIP Publishing purport or intend to grant permission to reuse material to which it does not hold appropriate rights.
   You may not alter or modify the Material in any manner. You may translate the Material into another language only if you have licensed translation rights. You may not use the Material for promotional purposes.
3. The foregoing license shall not take effect unless and until AIP Publishing or its agent, Copyright Clearance Center, receives the Payment in accordance with Copyright Clearance Center Billing and Payment Terms and Conditions, which are incorporated herein by reference.
4. AIP Publishing or Copyright Clearance Center may, within two business days of granting this license, revoke the license for any reason whatsoever, with a full refund payable to you. Should you violate the terms of this license at any time, AIP Publishing, or Copyright Clearance Center may revoke the license with no refund to you. Notice of such revocation will be made using the contact information provided by you. Failure to receive such notice will not nullify the revocation.
5. AIP Publishing makes no representations or warranties with respect to the Material. You agree to indemnify and hold harmless AIP Publishing, and their officers, directors, employees or agents from and against any and all claims arising out of your use of the Material other than as specifically authorized herein.
6. The permission granted herein is personal to you and is not transferable or assignable without the prior written permission of AIP Publishing. This license may not be amended except in a writing signed by the party to be charged.
7. If purchase orders, acknowledgments or check endorsements are issued on any forms containing terms and conditions which are inconsistent with these provisions, such inconsistent terms and conditions shall be of no force and effect. This document, including the CCC Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

This Agreement shall be governed by and construed in accordance with the laws of the State of New York. Both parties hereby submit to the jurisdiction of the courts of New York County for purposes of resolving any disputes that may arise hereunder.

V1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.