

<https://doi.org/10.1038/s41524-025-01677-4>

A neural master equation framework for multiscale modeling of molecular processes: application to atomic-scale plasma processes

Check for updates

Shoubhanik Nath¹, Joseph R. Vella^{2,4}, David B. Graves³ & Ali Mesbah¹ ✉

Plasma-surface interactions (PSI) play a crucial role in microelectronics fabrication; however, their multiscale nature and array of complex, often unknown interactions make computational modeling of PSIs extremely difficult. To this end, we propose a general neural master equation (NME) framework that uses master equations to describe the dynamics of a molecular process, wherein neural networks learned from atomistic simulations represent unknown transitions between different system states. By leveraging the physics-based structure of master equations and data-driven state transitions, the NME framework promotes generalizability and physics interpretability, and can bridge disparate length and time scales. The framework is demonstrated for multiscale modeling of Si atomic layer etching and reactive ion etching, where the learned NME-based surface kinetic models exhibit good predictive and extrapolative capabilities for predicting experimentally relevant observables as a function of process parameters. The NME-based surface kinetic models obey physical constraints, which are violated in models based on neural ordinary differential equations. The proposed NME framework for multiscale modeling of molecular processes can pave the way for the discovery of new chemistries and materials in atomic-scale plasma processes.

Interactions of low-temperature plasmas (LTP) with material surfaces are central to the fabrication of microelectronics^{1,2}, as approximately 30 percent of the semiconductor manufacturing chain relies on plasma processing. A holistic understanding of the interactions of plasma with materials, termed plasma-surface interactions (PSI), is essential for maintaining greater uniformity in microelectronic features, implementing atomic-scale control and reducing defects, as well as for exploring new chemistries, materials, and fabrication techniques^{3,4}. Despite huge strides in plasma modeling and surface science simulations, model-based investigations of atomic-scale plasma processes remain largely limited due to the lack of mechanistic models of complex PSI⁵.

The main challenges in modeling PSI stem from the vast array of possible atomistic processes at play. These include secondary electron emission⁶, spontaneous etching^{7,8}, surface reactions⁹, surface adsorption and modification¹⁰, surface charging¹¹, radical recombination¹⁰, diffusion¹², and knock-on collisions¹³, to name a few. Moreover, PSI strongly depend on the

plasma sheath, material surface condition, and bulk material properties, each of which is characterized by physics at different length and time scales⁵. Hence, PSI involve multiscale processes, spanning length scales from Å to cm and time scales from picoseconds to seconds. A comprehensive description of PSI must take into account the quantum interactions at the plasma-material interface, mesoscopic physics in the plasma sheath, and the macroscopic variations in the bulk plasma.

Multiscale modeling of PSI has received renewed interest as plasma processing of complex interfaces is becoming increasingly important in a wide range of emerging LTP applications⁴. Common multiscale modeling strategies for PSI include combining molecular dynamics (MD) and Monte-Carlo (MC) simulations^{5,12,14–16}, coupling kinetic MC (KMC) simulations of surface with fluid descriptions of the bulk plasma^{17,18}, rate equation approaches for describing specific interactions where kinetic rate constants are obtained from MD simulations¹⁹, and semiclassical models for charge transfer using separate Boltzmann equations for electrons and holes on

¹Department of Chemical and Biomolecular Engineering, University of California at Berkeley, 94720 Berkeley, CA, USA. ²U.S. Department of Energy National Laboratory, Princeton Plasma Physics Laboratory, 08540 Princeton, NJ, USA. ³Department of Chemical and Biological Engineering, Princeton University, Princeton, 08536 Princeton, NJ, USA. ⁴Present address: Tokyo Electron Ltd., 12205 Albany, NY, USA. ✉ e-mail: mesbah@berkeley.edu

either side of the interface along with quantum mechanical matching conditions at the interface²⁰. However, these multiscale modeling strategies are generally tailored to describe one or a subset of plasma-induced surface processes⁵, often ignoring the interplay between a complex array of interactions and thus limiting the ability to “transfer” the model to similar systems in a systematic and efficient manner. Additionally, these approaches rarely provide equation-oriented model representations that can readily capture the effects of process-level operating conditions and parameters, and can often be prohibitively expensive.

Interactions at the plasma-solid interface can also be affected by the subsurface, for example, when modified through ion bombardment¹⁵. Typical strategies to account for the subsurface include assuming a well-mixed zone (called the mixing layer) and using Arrhenius-type global reaction models to describe surface reactions^{21–24}. Such classical reaction rate models are empirical in nature and typically rely on experimentally determined reaction rate laws, where a key challenge stems from accounting for the complex series of reactions that can occur at the plasma-solid interface. On the other hand, KMC models describe the well-mixed subsurface layer of fixed depth through which chemical species can travel and reactions can progress^{5,15,16}. KMC models evolve probabilistic events through stochastic sampling. As such, they do not provide analytical kinetic equations and can be computationally expensive.

In recent years, machine learning-assisted approaches have been widely used for data-driven modeling of plasma physics, chemistry, and PSI (e.g.,^{25–32}). While there has been fairly significant progress on physics-informed, learning-assisted modeling of the bulk plasma behavior^{33–35}, much of the machine learning work related to PSI modeling has focused on creating black-box, surrogate models for plasma-induced surface effects^{36–38}. Although such black-box models can exhibit good predictive capabilities over the range of processing conditions covered in their training data, they do not generally provide interpretable representations of the multiscale processes governing PSI. Integration of physics into data-driven models can reduce overfitting, impart interpretability, extend extrapolative capabilities beyond the range of training data, and promote generalization to new systems^{39–41}. As such, fusing physics-based and data-driven modeling can be particularly useful for problems involving complex multiphysics.

Physics knowledge can be included into data-driven models either through the loss function, as in physics-informed neural networks⁴², or through composite equations composed of physics-based expressions, such as conservation laws, and function approximators. The latter approach is termed as universal differential equations (UDEs)⁴³, where parts of a differential equation are represented by neural networks or other function approximators as a substitute for unknown physics, such that the entire equation is differentiable⁴⁴. UDEs have emerged as a powerful tool in scientific computing for dynamic modeling of partially known systems⁴⁵, as well as discovery of governing equations of dynamical systems⁴⁶. Example applications include modeling of nucleation kinetics⁴⁷, neural systems in neuroscience⁴⁵, and pandemic outbreaks⁴⁸. The universality of UDEs, their ability to retain physics through the model structure, and the differentiable nature of the resulting equations can make them highly suited for inference of multiscale and multiphysics problems.

Master equations are widely used to model molecular processes^{49,50}, such as adsorption from gas to an adsorbent surface⁵¹, mRNA interactions with a promoter⁵², and particle hopping in space during diffusion⁵³. This work presents a neural master equation (NME) framework for multiscale modeling of PSI. The proposed NME framework builds a kinetic model through master equations, where unknown transition probabilities, which are mesoscopic averages of quantum-mechanical interactions, are represented with neural networks to form a set of UDEs. As such, the NME framework bridges the microscopic and mesoscopic scales, as it provides mesoscopic rate equations where the transition rates are obtained from atomistic simulations. Additionally, the proposed framework is able to admit spatial variations of quantities of interest that cannot be readily cast into a master equation form, such as those arising from transport.

The NME framework is demonstrated for multiscale modeling of Si atomic layer etch (ALE) with Cl₂ and Ar ion, as well as Si reactive ion etch (RIE) with F and Ar ion, two LTP processes that are widely used in semiconductor manufacturing^{54–56}. We demonstrate that the learned NME models provide an interpretable mesoscopic description of the evolution of surface processes by predicting experimentally relevant observables as a function of process parameters. The NME-based surface kinetic models exhibit extrapolative capabilities outside their training range in comparison to a fully data-driven model learned using neural ordinary differential equations⁵⁷. Our results suggest that the NME framework can be used as a viable physics-based surrogate for computationally expensive MD simulations to investigate PSI. The remainder of the paper is organized as follows. The section “General overview of the NME framework” presents a general overview of the NME framework. In the sections “NME-Based Surface Kinetic Model of ALE” and “NME-based surface kinetic model of RIE”, the NME framework is demonstrated on ALE and RIE, respectively. The section “Discussion” presents a discussion on the results and the broader use of the NME framework, followed by the details of the NME framework in the section “Methods”.

Results

General overview of the NME framework

In atomic-scale plasma processes such as RIE and ALE, there are multiple chemical elements that exist in different states, taking part in surface reactions and sputtering. For a molecular process with M species, with each species denoted by subscript m that undergoes $\mathcal{V}$ possible state transitions, the occupation probability of a species m is defined as the probability of occupying a state ν and is given by

$$P_m^\nu = \frac{N_m^\nu}{N_m}, \quad \forall \nu = 1, \dots, \mathcal{V}, \quad (1)$$

with P_m^ν denoting the occupation probability of species m in state ν , N_m denoting the total number of species m in all the $\mathcal{V}$ states, and N_m^ν denoting the number of species m in state ν . The dynamics of the state-to-state transitions can be represented by master equations⁵⁸

$$\frac{dP_m}{dt} = \mathbf{W}_m \mathbf{P}_m, \quad \forall m = 1, \dots, M, \quad (2)$$

where $\mathbf{W}_m \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ denotes the transition rates between the states of species m and $\mathbf{P}_m = [P_m^1, \dots, P_m^\mathcal{V}]^\top \in \mathbb{R}^{\mathcal{V} \times 1}$ is the vector of the occupation probabilities for species m in all of its $\mathcal{V}$ possible states.

Depending on the type of molecular interactions (e.g., quantum interactions for chemical reactions, or classical interactions for diffusion⁵⁹), determining the transition probabilities in Eq. (2) can be particularly challenging. Here, we use neural networks as universal approximators^{60,61} to learn the transition probabilities in Eq. (2) as a function of local inputs (e.g., incident ion energy and dose) to an atomic-scale process. As such, the master equation takes the form of a composite differential equation with neural network components, as given by

$$\frac{d\hat{\mathbf{P}}_m}{dt} = \hat{\mathbf{W}}_m(\mathbf{x}; \Theta) \hat{\mathbf{P}}_m, \quad \forall m = 1, \dots, M, \quad (3)$$

where $\hat{\mathbf{P}}_m = [\hat{P}_m^1, \dots, \hat{P}_m^\mathcal{V}]^\top$ is the vector of predicted occupation probabilities of species m and $\hat{\mathbf{W}}_m(\mathbf{x}; \Theta) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\Theta} \mapsto \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ is a neural model of the transition matrix, with $\mathbf{x}$ being the inputs to the system and Θ being the vector of learnable parameters.

In addition to state transitions, ion bombardment can result in incorporation of surface species into the bulk solid⁶² and ion-enhanced diffusion in the bulk solid^{63–69}. The ion-enhanced diffusion can result in the redistribution of material in the mixed layer and, thus, must be accounted for during state transitions, as during plasma etch, the moving etch front

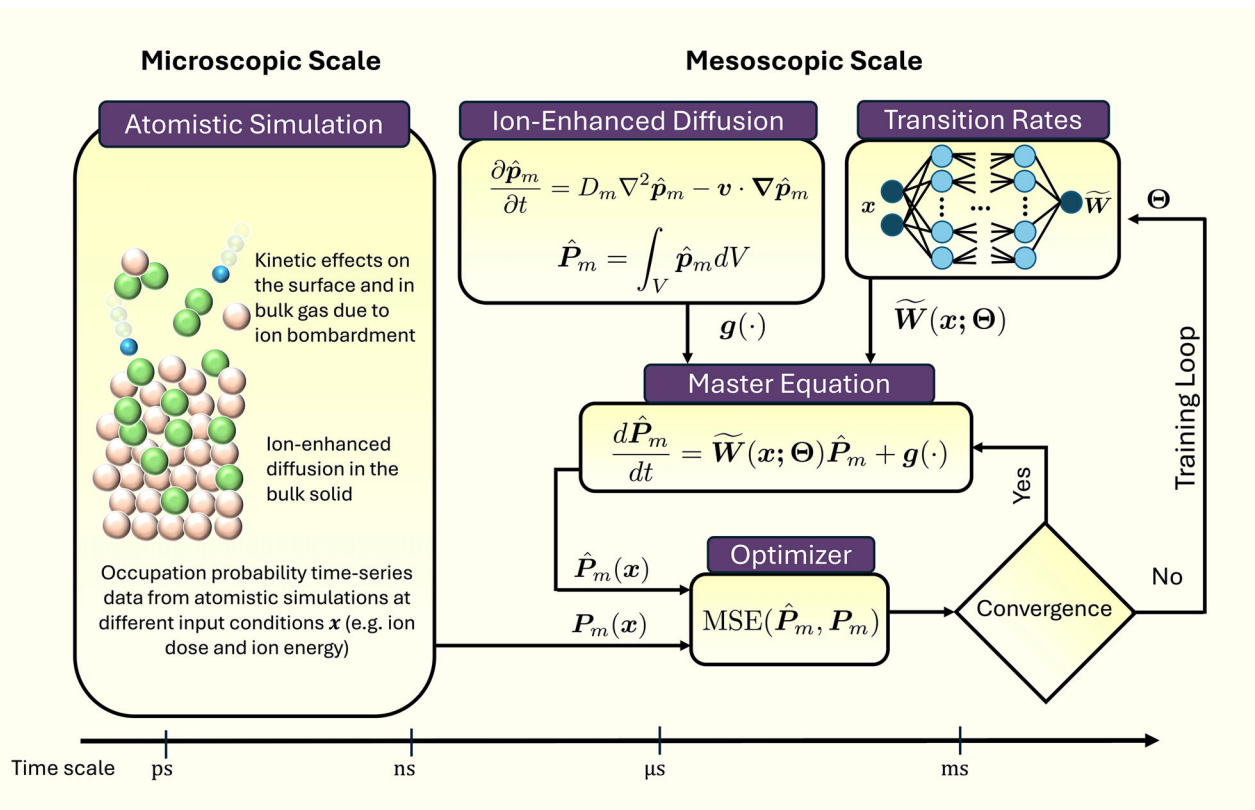


Fig. 1 | Overview of the neural master equation (NME) framework. At the microscopic scale, all-atom simulations are performed to generate time-series trajectory data of occupation probabilities $\{P_m\}_{m=1}^M$ for different values of system inputs x . The data are used to determine the learnable parameters of the neural model of

mesoscopic transition rates $\tilde{W}_m(x; \Theta)$ in the NMEs in Eq. (4) by minimizing the mean-squared-error loss function (Eq. (13)). The term $g(\cdot)$ is an addition to the master equation to account for possible probability currents due to additional physics, such as transport.

exposes underlying mixed layer material to the surface, effectively causing a state transition. The amount of material that is exposed depends on the local mixed layer concentration, which is affected by ion-enhanced diffusion. This non-linear state transition appears as an additive term to Eq. (3) as given by

$$\frac{d\hat{P}_m}{dt} = \tilde{W}_m(x; \Theta)\hat{P}_m + g(\cdot), \quad \forall m = 1, \dots, M, \quad (4)$$

where $g(\cdot)$ is any linear or non-linear contribution to the probability current due to transport or moving boundaries. Figure 1 provides an overview of the NME framework. Ensemble atomistic simulations at different input conditions of ion energy and fluence are used to generate time-series trajectory data of “ground truth” occupation probabilities $\{P_m\}_{m=1}^M$ to train the learnable parameters of $\tilde{W}_m(x; \Theta)$ in Eq. (4).

NME-based surface kinetic model of ALE

The NME framework was used to obtain a mesoscopic kinetic model of surface processes for the case of Si ALE with Cl_2 and Ar^+ . ALE is a layer-by-layer removal of material that progresses cyclically⁷⁰. During the first half cycle, a reactive gas is adsorbed on the surface of the material to be etched, whereas during the next half-cycle the modified surface is bombarded by ions to sputter material off the surface. The two steps occur sequentially and, hence, are temporally decoupled from each other.

In the ALE process at hand, Cl adsorption equilibrates quickly compared to the cycle time of adsorption, with Cl atoms fully covering all available active sites according to the Langmuir isotherm⁷¹. Hence, the dynamics of the adsorption process were ignored and, instead, an equilibrium coverage of Cl atoms was assumed at the start of every bombardment cycle. Three distinct states in which Cl can be present were considered, i.e.,

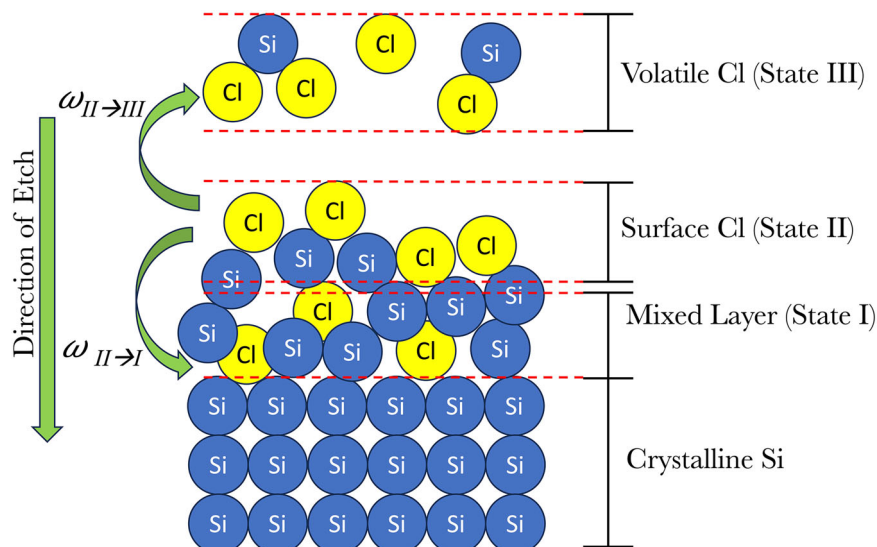
$v \in \{\text{I, II, III}\}$. Figure 2 depicts these states and the transitions among them. In state I, Cl atoms are mixed with bulk Si forming a mixed layer. This state is characterized by diffusion processes occurring in the mixed layer, with Cl only being able to interact with itself and Si. In state II, Cl atoms are present on the exposed surface of the Si-Cl block and are capable of interacting with both Ar^+ and Si. It was assumed that Ar^+ does not interact with Cl or Si in the mixed layer, and that the mixed layer composition can vary with depth, as opposed to the homogeneous mixed layer assumption in ref. 16. In state III, Cl atoms are present in bulk gas as Si_xCl_y . Due to high-energy Ar^+ bombardment, Cl atoms can transition from the surface (II) to the bulk gas phase (III). Additionally, Cl atoms can be pushed down from the surface into the subsurface mixed layer (I), where they mix with Si. Other transitions, such as recombination from the bulk gas phase to the surface, are also possible, but were ignored here for simplicity. The transition from state II to III can be broken down into a series of first-order transitions for different Cl species in the bulk (Si_xCl_y); however, all such transitions were lumped together.

Following Eq. (4), the set of NMEs for the three Cl states can be easily derived (See Supplementary Section 1.2). A value of $1 \times 10^{-19} \text{ m}^2/\text{s}$ was taken for the diffusion coefficient⁶⁶. Since the length and time scales change with time, the problem becomes stiff and, thus, an adaptive solver was used to solve the coupled, non-linear set of NMEs for the occupation probabilities.

Model validation and predictions

MD simulations were performed to generate high-resolution time series data for the occupation probability of Cl in the three states (I, II, and III) at different conditions of incident Ar^+ energy and dose, exactly following the parameters used in ref. 62. These time series data were used to train the surface kinetic model for Si ALE with Cl_2/Ar^+ . Only one ALE cycle was simulated, starting with a pristine Si block, covered with an equilibrium

Fig. 2 | Schematic representation of the three overarching states of Cl atoms in Si ALE with Cl_2 and Ar^+ . Blue circles represent Si atoms, while yellow circles represent Cl atoms. Cl atoms can be either mixed with Si in the subsurface mixed layer (state I), on the exposed surface (state II), or in the bulk gas as atomic or compound Cl (state III).



concentration of Cl atoms. In the MD simulations, sputtered species were taken out of the simulation box, and only species on the surface and in bulk Si were kept. One sample was taken per impact of Ar^+ ; thus, the total length of a time series at a particular condition is the number of ion impacts at that condition.

From ensemble MD simulations with three realizations per condition, time series data of the occupation probabilities were calculated at different conditions of Ar^+ energy and dose. Ion energies from 30 to 70 eV with intervals of 1 eV and ion dose between 100 and 1000 impacts at intervals of 50 impacts were considered, totaling 779 distinct time series to train the surface kinetic model. The range of ion flux for the above impacts was 3.14×10^{13} – 3.14×10^{14} ion/cm²-s, as is of the order seen in experiments⁷¹. The ion dose and ion energy were the only inputs $\mathbf{x}$ to the neural network $\widehat{W}_m(\mathbf{x}; \Theta)$ in Eq. (4). To keep parity with experimental results, the ion fluence in the MD simulations was selected to be the same as that in experiments⁶².

To classify Cl into the proposed states, the coordination number of all Cl atoms in the system was calculated: those having coordination number greater than two were classified as occupying state I, while all others were classified into state II, with the reasoning that Cl atoms inside the densely packed mixed layer would feel the effect of multiple neighboring species. This classification was based on heuristic arguments; however, more formal distinctions can be made, e.g., based on bond orders of individual Cl atoms from the modified reactive empirical bond-order potential used in ref. 72. The training details of NMEs in Eq. (4) for the ALE system are provided in Supplementary Section 2.1.

The transition rates in Eq. (4) were learned using the aforementioned time series data, as shown in Fig. 1. The resulting surface kinetic model was tested on combinations of ion dose and energy not present in the training dataset. Details of the neural network models, training procedure, and loss curves can be found in Supplementary Section 2.1. Figure 3 shows the comparison between the predicted occupation probability dynamics and those obtained from MD simulations; the statistical evaluation of the model performance is given in Table 1. The predicted occupation probability dynamics of all three states are in good quantitative agreement with test data for a variety of different ion dose and energy conditions. As bombardment starts with an equilibrium Cl coverage on a pristine Si block, the probability of finding Cl at the surface is unity. This probability decreases with time as Cl atoms transition to states I and III due to Ar^+ bombardment. However, exponential decay in the surface occupation probability is arrested by the non-linear transition of Cl atoms to the surface through the moving etch front, which exposes the underlying Cl atoms of the mixed layer.

The surface kinetic model retains the physical structure of the system by using the master equations to describe the state transitions and the ion-

enhanced diffusion for transport in the mixed layer. Consequently, it also uses relatively fewer trajectories (779) compared to learning a fully data-driven model³⁸, and ensures that mass conservation, Eq. (11), is always satisfied. The retention of the physical form of the dynamical equations using NMEs promotes model performance for a wide range of operating conditions of ion energy and dose, while providing physically consistent results.

Figure 4 shows the profile of Cl atomic fraction in the mixed layer at 1000 Ar^+ ion impacts. The profile arises due to competition between the rate of removal of material from the surface and rate of diffusion into the mixed layer. The penetration depth decreases with increasing ion energy. This is because at higher incident ion energies, Cl atoms are more rapidly sputtered off the surface due to higher etch rate, which results in subsurface Cl being exposed more rapidly. No clear trend between the total Cl concentration and the ion energy is observed due to the highly nonlinear nature of the process. At 40 eV, the ion energy is not sufficient to push too many Cl atoms into the mixed layer. The ability of Ar^+ to push Cl atoms into the mixed layer increases with increasing energy, as seen for 60 eV and 80 eV, where there are more Cl atoms in the mixed layer compared to 40 eV. For 100 eV, the etch front depletes fastest, and Cl atoms rapidly and nonlinearly transition from the mixed layer to the surface. This leads to fewer Cl atoms in the mixed layer and a smaller penetration depth. The total Cl and its penetration depend upon the competition between the etch and diffusive transport away from the etch front. The atomic fraction depth of 22 Å for 100 eV is in agreement with MD simulations from ref. 62, even though the surface kinetic model was not trained on atomic fraction data. To obtain the relevant concentrations from the occupation probabilities, it is sufficient to multiply the probabilities with the total number of Cl atoms in the system. The time evolution of number of Cl atoms is shown in Fig. 5. Since the bombardment step of ALE is considered to be a closed system with respect to Cl, the dynamics closely follow the time evolution of occupation probabilities.

Figure 6 shows the evolution of Cl uptake and Si etched over ALE cycles, and compares them with MD simulations from ref. 62. The representative first-order dynamics of Cl is shown only for the first half cycle. Cl adsorption process equilibrates very fast; hence, an equilibrium coverage of Cl was assumed at the start of every bombardment half cycle^{62,71}. The cases of 80 eV and 100 eV represent an extrapolation from the training data, since the training range was 30–70 eV. The surface kinetic model predicts incremental increases in the final Cl uptake from cycle to cycle (Fig. 6a), as increasingly more Cl atoms are incorporated into the mixed layer. More efficient removal of Cl is seen at higher ion energies in both the surface kinetic model predictions and the MD simulations.

Fig. 3 | Performance of the surface kinetic model for ALE on test data. The solid blue line, black dash-dot line, and red dashed line represent the occupation probabilities predicted by the surface kinetic model in the mixed layer (I), surface (II), and bulk gas (III), respectively, while the open circles denote the corresponding occupation probabilities from MD simulations. **a** 119 Ar⁺ impacts at 49 eV energy. **b** 261 Ar⁺ impacts at 51 eV energy. **c** 311 Ar⁺ impacts at 61 eV energy. **d** 593 Ar⁺ impacts at 43 eV energy. These combinations of ion dose and energy were not included in the training data, but the ion energy is within the training range of 30–70 eV.

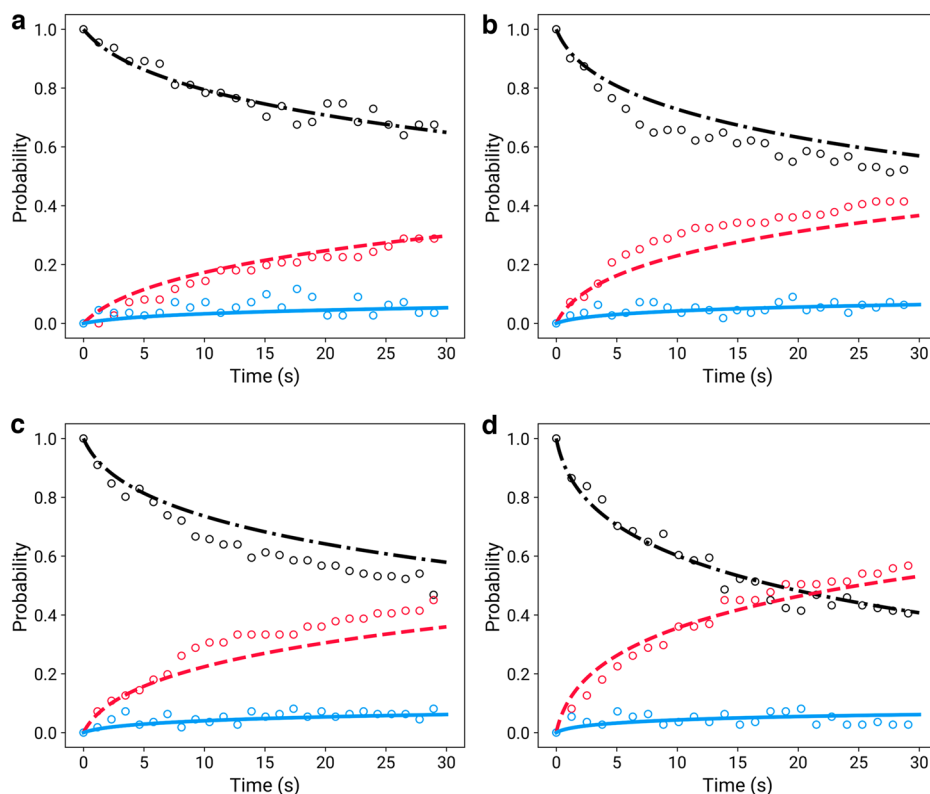


Table 1 | The root mean squared error (RMSE) between the surface kinetic model predictions and MD simulations for all four cases of Fig. 3

Statistics	Mixed Layer	Surface	Bulk Gas
RMSE	$0.024 \pm 5E-4$	$0.051 \pm 2E-3$	$0.046 \pm 2E-3$

The surface kinetic model correctly predicts that the Cl retained in the substrate post-bombardment increases across cycles. While the prediction from the surface kinetic model is of the same order as MD simulations, the predictions are off by 0.15–1 monolayer (ML). The primary cause for this mismatch can be the classification of Cl atoms into states. The classification was made on the basis of coordination number, which is a discrete classifying feature and, thus, is limited. The surface kinetic model overestimates the amount of Cl left on the surface post bombardment, compared to MD simulations, which predict almost no Cl at the surface post bombardment. This reduces the amount of Cl that can be absorbed in the subsequent adsorption half cycle. It should be noted that the amount of Cl left on the surface is a matter of classification. A more accurate representation of the physical system may be made through using bond orders and bond energies of the pair potential used in MD simulations⁷². However, the cut-off value for the bond energies and orders is a parameter that will require tuning from system to system. Other potential causes for the mismatch could be the inability of the surface kinetic model to accurately account for the increase in active sites due to ion bombardment⁷³ and the deficiency of the MD simulations in capturing diffusion, both of which can affect the concentration of Cl at the surface and subsurface. These effects can further cause a change in the total Cl uptake across cycles.

For the total Si etched (Fig. 6b), the model predicts an approximately constant etch per cycle. A very good agreement is found with the results of ref. 62, demonstrating the predictive capability of the NME-based surface kinetic model. The model can capture the physical idiosyncrasies of the system and maintain quantitative agreement, all while drastically reducing the problem to one of solving a set of ODEs.

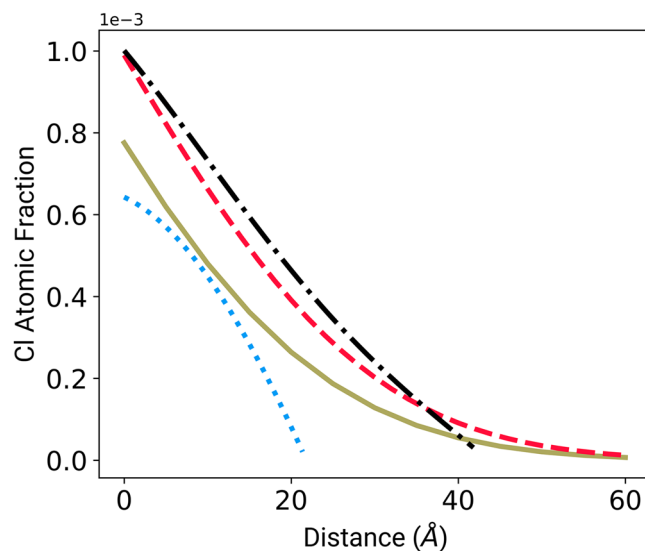


Fig. 4 | Predicted profile of the Cl atomic fraction in the mixed layer. Profiles are drawn at the end of 1000 Ar⁺ impacts as predicted by the surface kinetic model. The solid golden line is at 40 eV, the red dashed line at 60 eV, the black dash-dot line at 80 eV, and the blue dotted line at 100 eV.

A discrepancy is seen in the curvature gradient of the etch rate in Fig. 6b. However, when scalar values of the transition probabilities are judiciously chosen for particular conditions and the surface kinetic model is solved numerically with constant values of the transition probabilities, the correct curvature and a good fit is obtained, underlying the validity of the master equations. As neural networks are used to approximate the transition probabilities, the proposed approach is prone to getting stuck in local minima, and may not represent all of the physics correctly. Possible remedies for this are to include all the species and their respective states in the master equations and impose physically meaningful constraints on the

loss function. Additionally, the surface kinetic model does not show a cyclic steady state for Cl uptake, as seen in long-time MD simulations. Potential

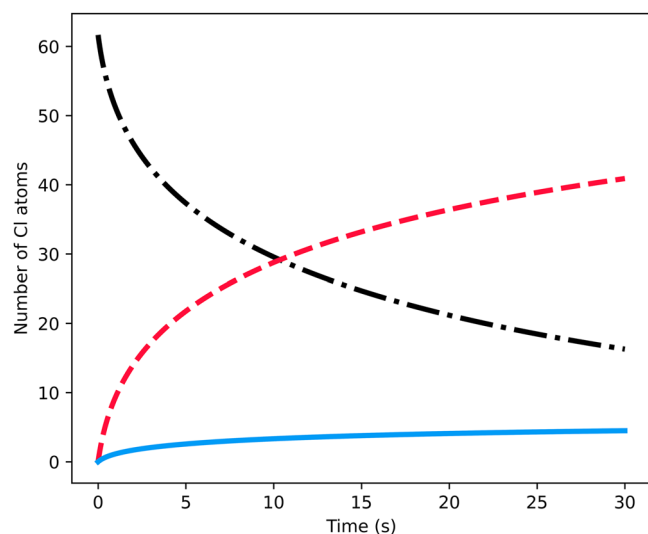


Fig. 5 | Time-evolution of the number of Cl atoms as predicted by the surface kinetic model for ALE for the three states under 1000 Ar^+ impacts at 30 eV. The solid blue line represents the mixed layer (I), the black dash-dot line represents the surface (II), and the red dashed line represents the bulk gas (III).

reasons for this could be using a constant diffusion coefficient, or the absence of depth-dependent diffusion⁶⁷.

The performance of the proposed NME framework is compared to that of a NODE model of the form of Eq. (14). To ensure that the results are physical, the loss function is constrained (See Supplementary Section 3) to obey Eq. (11) and provide probabilities in the interval [0,1]; this approach is termed as fully constrained NODE (FC-NODE). Figure 7 shows the performance of FC-NODE against test data. FC-NODE performs poorly on test data and both constraints, although not violated during training, are violated on test data. Unlike NME (Eq. (4)), FC-NODE (Eq. (14)) does not retain the physics-based structure of state transitions or transport. Furthermore, while Eq. (11) is built into the structure of NMEs, it only appears as a soft constraint in the loss function during training of the FC-NODE.

NME-based surface kinetic model of RIE

Surface kinetic modeling of RIE of Si with F and Ar^+ is also studied using the NME framework shown in Fig. 1. RIE is a continuous process, where a plasma chamber is filled with reactive and neutral gases. Accelerated ions sputter the surface, and reactive neutrals and radicals chemically react with the exposed surface⁷⁴. As in the Si ALE system above, three states were considered to represent the RIE system at hand: F atoms can exist in the mixed layer (state I), at the exposed surface of Si (state II), and as volatiles in the gas phase (state III). The NME model structure was kept the same as that for ALE, with two distinct differences: (i) RIE is an open system with respect to F atoms, and (ii) there is an additional adsorption process for F on the surface of Si. Hence, the master equations were modified to account for these differences; derivations are provided in Supplementary Section 1.3.

Fig. 6 | Predictions of the surface kinetic model for ALE cycles at 1000 Ar^+ impacts, compared with MD simulations from ref. 62. An equilibrium Cl coverage was assumed at the start of each bombardment cycle since Cl adsorption equilibrates very fast⁷¹. The blue, golden, red, and black lines represent Ar^+ energies of 40 eV, 60 eV, 80 eV, and 100 eV, respectively. **a** Cl uptake in monolayers (ML); the adsorption half cycle is shown only for the first cycle (dashed line in teal). The dashed lines correspond to MD simulations, while the solid lines are the surface kinetic model predictions. **b** Distance of Si etched (Å). The open circles correspond to MD simulations, while the solid lines are predictions from the surface kinetic model.

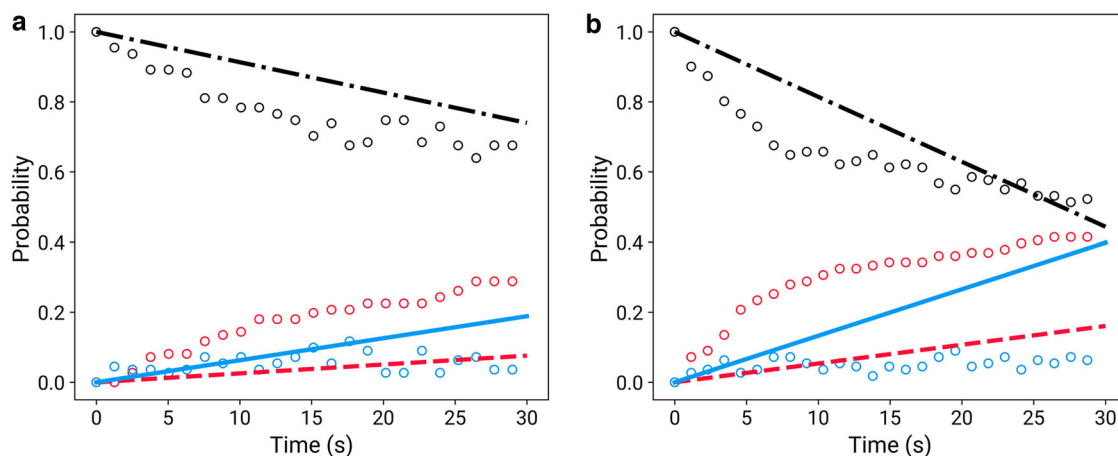
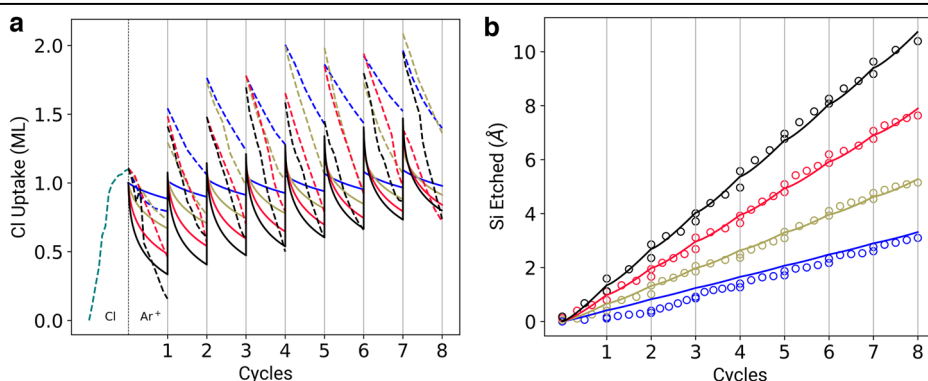
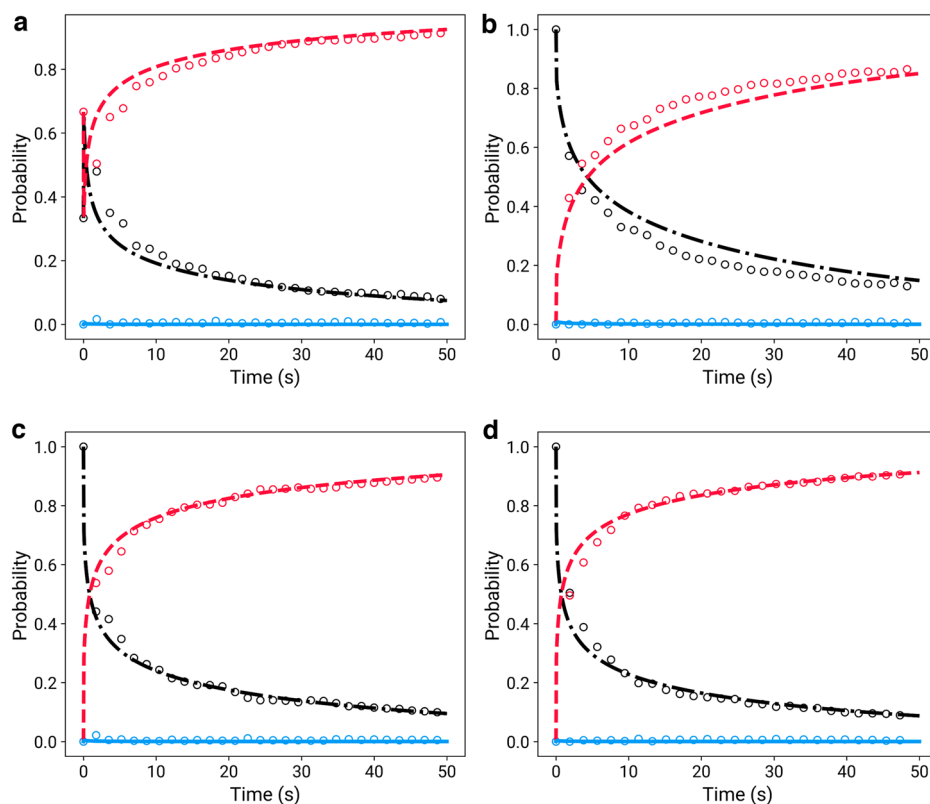


Fig. 7 | Performance of the FC-NODE model on test data. The lines denote predictions from the model, while the open circles denote the test data from MD simulations. The solid blue line, the black dash-dot line, and the red dashed line

represent the mixed layer (I), the surface (II), and the bulk gas (III) occupation probability dynamics, respectively. **a** 119 Ar^+ impacts at 49 eV energy. **b** 261 Ar^+ impacts at 51 eV energy.

Fig. 8 | Performance of the surface kinetic model for RIE on test data. The lines denote predictions from the model, while the open circles denote test data from MD simulations. The solid blue line, the black dashed-dot line, and the red dashed line represent the mixed layer (I), the surface (II), and the bulk gas (III) occupation probability dynamics, respectively. **a** 1100 total impacts at 90 eV Ar^+ energy. **b** 560 total impacts 51 eV Ar^+ energy. **c** 864 total impacts 79 eV Ar^+ energy, and **d** 952 total impacts 29 eV Ar^+ energy.



Similar to the Si ALE system, the characteristic length and time scales of ion-enhanced diffusion within bulk Si are $D/v(t)$ and $D/v(t)^2$, respectively. The atomic fraction is used, instead of the probability density function, to obtain a simpler representation of diffusion within bulk Si; however, the atomic fraction and probability density function share a linear relationship given by Eq. (9), where the total number of atoms N_m is now a function of time. There are two fluxes to the exposed surface: the flux of incident Ar^+ , J_{Ar} , and the flux of incident F atoms, J_{F} . The net rate of addition of F into the system is described by an additional ODE that can be solved independently from the set of NMEs. In an actual RIE system, gases are continuously pumped out from the etch chamber. Here, it is assumed that any gases taken out are also in state III, an assumption that can be dropped by adding a removal term to the surface kinetic model. Another difference from the Si ALE system is the rate of adsorption of F to the exposed surface.

Model validation and predictions

MD simulations with appropriate interatomic potentials are extensively used to study surface dynamics in RIE processes^{13,75–78}. Following ref. 79, we used MD simulations to generate time series data of F occupation probabilities to train the NME-based surface kinetic model for RIE of Si.

The architecture of the neural networks representing the transition rates in Si RIE was kept the same as that for Si ALE, as similar kinetics are expected in both systems. The transition rates are functions of the incident ion energy and the combined dose of F and Ar in the system. The flux ratio of F to Ar was kept constant at 5. Ion energy was varied from 20 eV to 80 eV at intervals of 2 eV, while the combined dose values were in the range of 500 to 1000 impacts with an interval of 25 impacts, resulting in 651 distinct combinations. Figure 8 shows the performance of the NME-based surface kinetic model on unseen test data, where neither the dose nor the ion energy is present in the training dataset. Furthermore, a condition of 1100 total impacts at 90 eV Ar^+ energy, outside of the domain of training data, is also considered. As can be seen, the NME-based surface kinetic model exhibits very good agreement with test data and can also generalize beyond the training dataset, as evident from Fig. 8a. The mixed layer concentration is

negligible since the extremely strong Si-F bond⁸⁰ significantly weakens the Si-Si bond and prevents F from being pushed into the mixed layer.

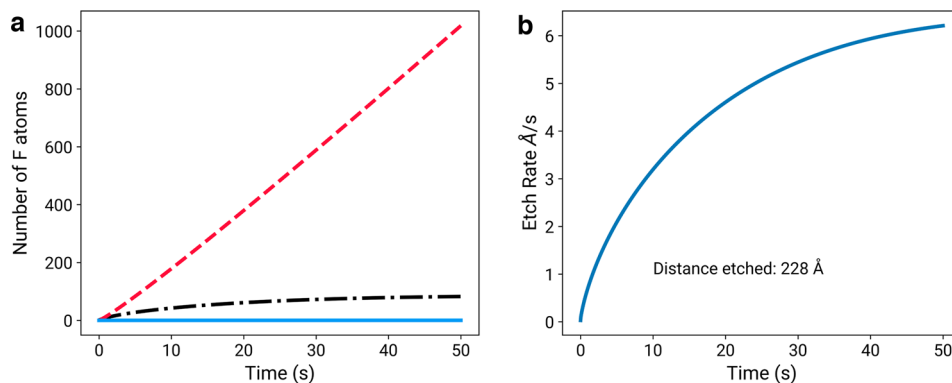
Figure 9a shows the F occupation dynamics at different states with the learned model for a dose of 1100 total impacts and ion energy of 90 eV. The number of F atoms in the bulk gas phase quasi-linearly increases due to the flux of F atoms into the system, while that on the surface, plateaus and reaches a steady state due to competing sputtering and adsorption processes. Since the surface kinetic model also solves for the instantaneous velocity of the etch front, the instantaneous etch rate can be calculated. Figure 9b shows the dynamics of the instantaneous etch rate, and the area under the curve gives the total amount etched. The etch yield, which is the ratio of Si sputtered per incident Ar^+ , was not calculated here, but it can be modeled by including the master equation for Si.

Discussion

In this paper, an NME framework is introduced for multiscale modeling of plasma-surface interactions in atomic-scale plasma processes. The framework can be adopted for any system where chemical master equations can be used to describe the underpinning molecular processes, for example, in spin dynamics, lasers⁸¹, electroporation and electropore-transport⁸², and electron exchange from electrolyte phase to electrode surface⁸³, amongst others. Here, the NME framework is used to derive surface kinetic models for ALE and RIE of Si with Cl_2/Ar^+ and F/Ar^+ , respectively, as these systems are industrially relevant. The NME-based surface kinetic model is informed by the physics of state transitions of the system^{81–83}. The NME structure is adaptable, whereby additional physics can be accounted for.

The examples demonstrated for plasma processing have numerous reaction steps. Ideally, each reaction step should be modeled as a master equation in order to satisfy microscopic reversibility. The proposed NME framework can be readily used for master equations with microscopic reversibility, as the model structure makes no assumptions and proper NME training can result in a detailed balance being satisfied. However, for practical purposes, accounting for all the reaction steps and intermediate states

Fig. 9 | Predictions of the surface kinetic model for RIE. **a** Time-evolution of the number of F atoms at 1100 total impacts and 90 eV Ar⁺ energy. The solid blue line, black dash-dot line, and red dashed line represent the number of F atoms in the mixed layer (I), surface (II), and bulk gas (III), respectively. **b** Time-evolution of Si etch rate (Å/s) at 1100 total impacts and 90 eV Ar⁺ energy.



can become intractable. While full knowledge of intermediate states and transitions and their incorporation into the master equation preserve the detailed balance, one can lump intermediate states into an overall effective state. Omitting the states results in a set of reduced-order “quasi” master equations that yield a surface kinetic model useful for longer time scale investigation of the system of interest.

Retention of the physics of state transitions in the NME framework ensures that relevant constraints are inherently built into the structure of the model. Hence, predictions from the NME-based surface kinetic model do not show non-physical results, such as negative probabilities or sum of probabilities be greater than one at any time instant for any condition, as opposed to the black-box NODE models. Additionally, NME models whose transition rates are only parametrized by neural networks are expected to use lesser data compared to NODEs that use neural networks to represent the entire state transition dynamics⁴⁷. With the same amount of data, NME vastly outperforms FC-NODE, although FC-NODE may provide improved predictions when trained with substantially more data. However, FC-NODE does not guarantee mass conservation, unlike the NME model, as seen in Fig. 7.

Due to their physics-based structure, NME-based surface kinetic models are also capable of extrapolation to regimes beyond their training data. Another advantage of NME-based surface kinetic models is that they enable exploration of longer time scales than is possible by MD simulations. For example, the MD simulations in Fig. 6 done on one core of Intel Xeon Gold 6330 took on the order of 8–10 hours, while the surface kinetic model simulations done on one core of Apple M2 took approximately 5 minutes, representing a 99% decrease in computation time. The significant computational speed up provided by NME-based surface kinetic models makes comparisons between model predictions of system observables, like concentrations in different phases, and experimental observations, such as optical emission spectroscopy signals⁸⁴, possible. NME-based models can be deployed simultaneously with experiments or during fabrication, and be used to make online decisions. Furthermore, these models allow for exploration of large surfaces at longer length scales with ion energy distribution functions obtained from plasma simulations. This can be used in surface profile evolution with much smaller computational time, as opposed to traditional voxel-based methods with KMC⁸⁵. A possible application is the study of roughness and critical dimension uniformity over a wafer surface, or in a smaller feature.

A possible extension of the proposed surface kinetic models for the plasma etch processes is to learn the ion-enhanced diffusion coefficient, in lieu of using guess values (Eq. (8)). Atomic fraction data from MD simulations can be used to learn the effect of ion dose and energy on the diffusion coefficient and, thus, obtain more accurate values of the ion-enhanced diffusion coefficient. This would enable better predictions of penetration depth and concentration profiles in the mixed layer. The microscopic resolution of state transitions provided by the NME framework can be used in the discovery of new materials and chemistries in atomic-scale plasma processing. The resultant scale-bridging surface

kinetic models can also be used for surface evolution studies and recipe design for the next-generation semiconductor device fabrication.

Methods

Master equation for molecular processes

In molecular processes, the occupation probability for each species can be defined by an objective probability as given in Eq. (1), with the index m used for any species. The total possible states are fixed for a particular system with defined chemistry. The state-to-state transitions can be modeled as first-order processes that occur with some transition rate, also known as transition probability. While the assumption of first-order transitions may not always be true, higher or fractional order transitions can be converted to pseudo first-order transitions. In essence, the state-to-state transitions describe the dynamics of a molecular process, wherein the occupation number of different states can be averaged to obtain approximate dynamical rate equations.

For a molecular process consisting of M species that undergoes $\mathcal{V}$ possible state transitions among all its species Eq. (2) constitutes a set of linear ordinary differential equations (ODEs), describing the rate of change of occupation probabilities $\{P_m\}_{m=1}^M$. The rate of change in the occupation probability P_m^v due to transitions into and away from any state v is given by

$$\frac{dP_m^v}{dt} = \sum_{\substack{v' \neq v \\ v'=1}}^{\mathcal{V}} \{-\omega_m^{v'v} P_m^v + \omega_m^{vv'} P_m^{v'}\}, \quad \forall v = 1, \dots, \mathcal{V}, \quad \forall m = 1, \dots, M. \quad (5)$$

The first term in Eq. (5) represents all possible transitions from state v with an associated transition rate of $\omega_m^{v'v}$, while the second term represents all possible transitions to the state v from other states with an associated transition rate of $\omega_m^{vv'}$. Hence, the probability of transition from state v to state v' in a small time Δt is $\omega_m^{vv'} \Delta t$. These transition rates form the elements of the transition matrix $\mathbf{W}_m$ in Eq. (5).

Neural representation of transition probabilities

In PSI, the transition probabilities are mesoscopic averages of the many-body quantum interactions between different species. Density functional theory and ab-initio MD simulations have been widely used to study these quantum mechanical interactions for a variety of systems^{86–89}. While information of the energy barriers and other physical parameters would vastly improve the predictive power of mesoscopic reaction rate models⁹⁰, obtaining exact transition probabilities from many-body quantum interactions is often not possible.

Machine learning-based approaches have been developed to learn approximate representations for transition rates^{91–93} for chemical reaction networks approximated as Markov processes on continuous state space. However, these use simulation trajectories to model transition kernels, which take in the current state and output the state at the next time interval⁹¹, which make the kernel time-dependent, without an explicit dependence on

local system conditions. To mitigate that, we use neural networks to learn the transition probabilities in Eq. (5) as a function of system inputs (such as incident ion energy and ion dose), as opposed to the occupation probability at any time t . Thus, the neural networks represent discrete values of the transition rates, and not distributions, as has been done in the previous works^{91–93}. The resulting master equation with neural network components is given by

$$\frac{d\hat{P}_m^\nu}{dt} = \sum_{\nu' \neq \nu}^{\mathcal{V}} \left\{ -\tilde{\omega}_m^{\nu\nu'}(\mathbf{x}; \boldsymbol{\theta}) \hat{P}_m^\nu + \tilde{\omega}_m^{\nu'\nu}(\mathbf{x}; \boldsymbol{\theta}) \hat{P}_m^{\nu'} \right\}, \quad (6)$$

$$\forall \nu = 1, \dots, \mathcal{V}, \quad \forall m = 1, \dots, M,$$

where $\hat{P}_m^\nu$ is an approximation of the occupation probability P_m^ν , $\tilde{\omega}_m^{\nu\nu'}(\mathbf{x}; \boldsymbol{\theta})$ and $\tilde{\omega}_m^{\nu'\nu}(\mathbf{x}; \boldsymbol{\theta})$ denote (deep) neural networks that are a function of n_x system inputs, $\mathbf{x} \in \mathbb{R}^{n_x}$, and are parameterized by learnable parameters $\boldsymbol{\theta}$. We refer to the composite differential equation in Eq. (6) as a neural master equation (NME). The key advantage of NME is the flexibility that it offers in approximating the unknown transitions $\tilde{\omega}$ from atomistic simulation data while preserving the physics-based structure of (5).

Following Eq. (2), the set of NMEs for a molecular process with M species can be cast as

$$\frac{d\hat{\mathbf{P}}_m}{dt} = \tilde{\mathbf{W}}_m(\mathbf{x}; \boldsymbol{\theta}) \hat{\mathbf{P}}_m, \quad \forall m = 1, \dots, M, \quad (7)$$

where $\hat{\mathbf{P}}_m = [\hat{P}_m^1, \dots, \hat{P}_m^{\mathcal{V}}]^\top$ is the vector of predicted occupation probabilities of species m and $\tilde{\mathbf{W}}_m(\mathbf{x}; \boldsymbol{\theta}) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \mapsto \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ is a neural model of transition rates, whose elements consist of neural networks $\tilde{\omega}(\mathbf{x}; \boldsymbol{\theta})$ in Eq. (6). The parameters $\boldsymbol{\theta}$ constitute the concatenated vector of learnable parameters $\boldsymbol{\theta}$. Equation (7) is, in fact, a set of continuous-time UDEs⁴³ that synergizes the interpretable structure of the master equation with data-driven descriptions of hard-to-model transition rates. Equation 7 can be readily solved using state-of-the-art ODE solvers to obtain time-series predictions of occupation probabilities for different values of system inputs $\mathbf{x}$.

Ion-enhanced transport in bulk solid

The set of NMEs in Eq. 7 describes PSI consisting of reactions and sputtering events. However, ion bombardment has two other consequences: incorporation of surface species into the bulk solid⁶² and ion-enhanced diffusion in the bulk solid^{63–69}. Ion-enhanced diffusion is of great importance in small features, where characteristic diffusive length scales can become comparable to the feature size⁹⁴. Incorporation of surface species can be considered as a state transition from the surface to the bulk solid, thus lending itself to a master equation representation. However, diffusion-driven species transport due to ion bombardment gives rise to spatial variations in the probability density of the corresponding occupation probability⁹⁵ within a subsurface, amorphous mixed layer^{62,96–98}. It is important to account for the mixed layer in the description of PSI, particularly in plasma etch wherein the removal of material from the surface exposes the mixed layer. Additionally, the removal or addition of material, in plasma etch or deposition, respectively, results in a moving surface, which constitutes a convective transport phenomenon. This drift-diffusion transport in the mixed layer results in a non-linear state transition.

While diffusive processes can be described by master equations⁹⁹, the number of states to be accounted for can become prohibitively large. Instead, one can observe the process at timescales longer than the timescale of microscopic fluctuations to obtain a continuum diffusion equation within the bulk solid. These timescales are not short enough to resolve all the states and transitions associated with sputtering and reaction events on the surface. However, a master equation representation can be formulated for longer time state transitions on the surface by ignoring intermediate short-lived states. A consequence of this is that microscopic reversibility is no

longer applicable to the state transitions. This simplifying assumption of using only long timescale states and transitions allows significant reduction in the number of states to be considered in the proposed NME framework, and makes it viable for modeling PSI that involve ion-enhanced diffusion in the bulk solid. Diffusion at the atomic scale can be understood by considering the motion of atoms that enter the mixed layer with some momentum. These atoms undergo a series of random collisions with other atoms in the mixed layer. The position and momentum after collision depend only on the last collision, and the atoms do not retain memory of previous collisions, a hallmark of a Markov process. Hence, we describe the drift-diffusion transport in the mixed layer by

$$\frac{\partial \hat{\mathbf{P}}_m}{\partial t} = \mathbf{D}_m \nabla^2 \hat{\mathbf{P}}_m - \mathbf{v}(t) \cdot \nabla \hat{\mathbf{P}}_m, \quad (8)$$

where $\hat{\mathbf{P}}_m = [\hat{P}_m^1, \dots, \hat{P}_m^{\mathcal{V}}]^\top$ is the probability density of the corresponding occupation probability, $\hat{\mathbf{P}}_m$. Eq. (8) is in fact a Fokker-Planck equation, where the drift term is due to a moving boundary, instead of an externally applied field, and the diffusive term is assumed to be a constant. $\mathbf{D}_m$ is a diagonal matrix of the ion-enhanced diffusion coefficient of species m in the corresponding state ν , and $\mathbf{v}(t)$ is the velocity of the translating plasma-solid interface. The occupation probability in each state is

$$\hat{P}_m^\nu(t) = \int_{V^\nu} \hat{P}_m^\nu(\mathbf{x}, t) dV^\nu,$$

where V^ν is the hypervolume of state ν . The probability density is related to the atomic fraction by

$$\hat{\mathbf{P}}_m = \frac{\mathbf{n} \circ \mathbf{c}_m}{N_m}, \quad (9)$$

where $\circ$ is the Hadamard product, $\mathbf{n}$ is the vector of atomic densities of all states, assumed constant⁶⁶, N_m is the total number of species m in all the $\mathcal{V}$ states, and $\mathbf{c}_m$ is the vector of atomic fraction of species m in all states. Accordingly, the set of NMEs in Eq. (7) can be modified as

$$\frac{d\hat{P}_m^\nu}{dt} = \sum_{\nu' \neq \nu}^{\mathcal{V}} \left\{ -\omega_m^{\nu\nu'} \hat{P}_m^\nu + \omega_m^{\nu'\nu} \hat{P}_m^{\nu'} \right\} + \delta_{\nu\nu_s} A \hat{P}_m^{\nu_s} \mathbf{v}(t), \quad (10)$$

$$\forall \nu = 1, \dots, \mathcal{V}, \quad \forall m = 1, \dots, M,$$

where ν_s refers to the exposed surface state, $\delta_{\nu\nu_s}$ is the Kronecker delta function, and A is the area of the exposed surface. The final term in Eq. (10) represents a probability current for state ν , rendering the equations non-linear. Notice that Eq. (10) must follow mass conservation, which is equivalent to the summability relation of probabilities, i.e.,

$$\sum_\nu \frac{d\hat{P}_m^\nu}{dt} = 0. \quad (11)$$

To close the system of equations, boundary conditions are needed for Eq. (8). Probability current continuity must be imposed at the boundary between different states across which transitions can occur, while the current continuity equations must be consistent with mass conservation in Eq. (11). Conversely, substitution of Eq. (10) in Eq. (11) yields another boundary condition that ensures mass conservation.

Other transport phenomena that cause probability currents out of a state can be included in a similar manner, and will only contribute additional probability current terms in Eq. (10). Accordingly, the set of NMEs in Eq. 7

takes the form of Eq. (4)

$$\frac{d\hat{\mathbf{P}}_m}{dt} = \widetilde{\mathbf{W}}_m(\mathbf{x}; \Theta) \hat{\mathbf{P}}_m + \mathbf{g}(\cdot), \quad \forall m = 1, \dots, M, \quad (12)$$

where $\mathbf{g}(\cdot)$ is any linear or non-linear contribution to the probability current due to transport or moving boundaries, as described in Eq. (10).

Learning framework

The proposed NME framework is shown in Fig. 1. The training data for learning the neural model of transition matrix $\widetilde{\mathbf{W}}_m$ in Eq. 12 are obtained from atomistic simulations for different values of system inputs $\mathbf{x}$ (e.g., incident ion energy and ion dose in atomic-scale plasma processes). An ensemble of atomistic simulation data collected at different input values is used to generate time-series trajectory data of “ground truth” occupation probabilities $\{\mathbf{P}_m\}_{m=1}^M$. Accordingly, the learnable parameters Θ of the transition matrix $\widetilde{\mathbf{W}}_m(\mathbf{x}; \Theta)$, are determined by minimizing the mean-squared-error (MSE) loss

$$\mathcal{L}(\Theta) = \frac{1}{N} \sum_{i=1}^N \sum_{m=1}^M (\mathbf{P}_m(\mathbf{x}_i) - \hat{\mathbf{P}}_m(\mathbf{x}_i; \Theta))^2. \quad (13)$$

Here i denotes the trajectory index of time-series occupation probabilities for different values of inputs $\mathbf{x}_i$, with N being the total number of trajectories. The occupation probabilities $\{\hat{\mathbf{P}}_m(\mathbf{x}_i; \Theta)\}_{m=1}^M$ are predicted by numerical integration of the NMEs in Eq. 12 using standard ODE solvers (see Supplementary Section 2.1). Note that the differentiable nature of the NMEs allows for the use of backpropagation through the ODE solver when minimizing the loss function in Eq. (13).

Here, we briefly contrast the proposed NMEs with neural ordinary differential equations (NODEs)⁵⁷, which can be used to take a fully black-box approach to describing the time-evolution of occupation probabilities

$$\frac{d\hat{\mathbf{P}}_m}{dt} = \mathbf{f}(\hat{\mathbf{P}}_m, \mathbf{x}; \boldsymbol{\gamma}), \quad \forall m = 1, \dots, M, \quad (14)$$

where the function $\mathbf{f}$ is treated as a black box, approximated using (deep) neural networks with learnable parameters $\boldsymbol{\gamma}$. The NODEs formulation in Eq. (14) does not include any physics in the structure of $\mathbf{f}$ and, thus, does not provide an interpretable representation of the mesoscopic behavior of the molecular process. Yet, from a computational standpoint, a significant advantage of NODEs over NMEs is that NODEs can make use of the adjoint sensitivity method⁵⁷ to efficiently backpropagate through $\mathbf{f}$. The savings in computational cost are, however, offset by the poor predictive capabilities of NODEs and its non-physical predictions outside the training range, as demonstrated in the section “NME-Based Surface Kinetic Model of ALE”.

Data availability

The data is available from the authors upon reasonable request.

Code availability

The codes can be found in the Github repository <https://github.com/Mesbah-Lab-UCB/NODE-Training-for-Plasma-Surface-Kinetic-Model>.

Received: 26 December 2024; Accepted: 26 May 2025;

Published online: 15 July 2025

References

- National Academies of Sciences, M., Engineering et al. *Plasma Science: Enabling Technology, Sustainability, Security, and Exploration* (The National Academic Press, 2020).
- Oehrlein, G. S. et al. Future of plasma etching for microelectronics: challenges and opportunities. *J. Vac. Sci. Technol. B* **42** (2024).
- Graves, D. B. et al. Science challenges and research opportunities for plasma applications in microelectronics. *J. Vac. Sci. Technol. B* **42** (2024).
- Adamovich, I. et al. The 2022 plasma roadmap: low temperature plasma science and technology. *J. Phys. D: Appl. Phys.* **55**, 373001 (2022).
- Bonitz, M. et al. Towards an integrated modeling of the plasma-solid interface. *Front. Chem. Sci. Eng.* **13**, 201–237 (2019).
- Sharma, G. et al. Secondary electron emission and collisional effects in a two-electron temperature plasma sheath. *Contrib. Plasma Phys.* **63**, e202300020 (2023).
- Kanarik, K. J. et al. Predicting synergy in atomic layer etching. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **35**, 05C302 (2017).
- Kanarik, K. J., Tan, S. & Gottscho, R. A. Atomic layer etching: rethinking the art of etch. *J. Phys. Chem. Lett.* **9**, 4814–4821 (2018).
- Winters, H. F., Coburn, J. & Chuang, T. Surface processes in plasma-assisted etching environments. *J. Vac. Sci. Technol. B: Microelectron. Process. Phenom.* **1**, 469–480 (1983).
- Arts, K. et al. Foundations of atomic-level plasma processing in nanoelectronics. *Plasma Sources Sci. Technol.* **31**, 103002 (2022).
- Matsui, J., Nakano, N., Petrović, Z. L. & Makabe, T. The effect of topographical local charging on the etching of deep-submicron structures in SiO₂ as a function of aspect ratio. *Appl. Phys. Lett.* **78**, 883–885 (2001).
- Bogaerts, A. et al. Modeling of the plasma chemistry and plasma-surface interactions in reactive plasmas. *Pure Appl. Chem.* **82**, 1283–1299 (2010).
- Tinacba, E. J. C., Isobe, M. & Hamaguchi, S. Surface damage formation during atomic layer etching of silicon with chlorine adsorption. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **39**, 042603 (2021).
- Becker, M. & Sierka, M. Atomistic simulations of plasma-enhanced atomic layer deposition. *Materials* **12**, 2605 (2019).
- Guo, W., Bai, B. & Sawin, H. H. Mixing-layer kinetics model for plasma etching and the cellular realization in three-dimensional profile simulator. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **27**, 388–403 (2009).
- Kwon, O. & Sawin, H. H. Surface kinetics modeling of silicon and silicon oxide plasma etching. II. Plasma etching surface kinetics modeling using translating mixed-layer representation. *J. Vac. Sci. Technol. A* **24**, 1914–1919 (2006).
- Vanraes, P., Parayil Venugopalan, S. & Bogaerts, A. Multiscale modeling of plasma-surface interaction-general picture and a case study of Si and SiO₂ etching by fluorocarbon-based plasmas. *Appl. Phys. Rev.* **8** (2021).
- Zhang, D. & Kushner, M. J. Investigations of surface reactions during C2F6 plasma etching of SiO₂ with equipment and feature scale models. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **19**, 524–538 (2001).
- Filinov, A., Bonitz, M. & Loffhagen, D. Microscopic modeling of gas-surface scattering. I. A combined molecular dynamics-rate equation approach. *Plasma Sources Sci. Technol.* **27**, 064003 (2018).
- Bronold, F. X. & Fehske, H. Kinetic modeling of the electronic response of a dielectric plasma-facing solid. *J. Phys. D: Appl. Phys.* **50**, 294003 (2017).
- Gray, D. C., Tepermeister, I. & Sawin, H. H. Phenomenological modeling of ion-enhanced surface kinetics in fluorine-based plasma etching. *J. Vac. Sci. Technol. B: Microelectron. Nanometer Struct. Process. Meas. Phenom.* **11**, 1243–1257 (1993).
- Chang, J. P., Arnold, J. C., Zau, G. C., Shin, H.-S. & Sawin, H. H. Kinetic study of low energy argon ion-enhanced plasma etching of polysilicon with atomic/molecular chlorine. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **15**, 1853–1863 (1997).
- Marinov, D. Kinetic monte carlo simulations of plasma-surface reactions on heterogeneous surfaces. *Front. Chem. Sci. Eng.* **13**, 815–822 (2019).
- Vella, J. R., Hao, Q., Elgarhy, M. A., Donnelly, V. M. & Graves, D. B. A transient site balance model for atomic layer etching. *Plasma Sources Sci. Technol.* **33**, 075009 (2024).

25. Trieschmann, J., Vialotto, L. & Gergs, T. Machine learning for advancing low-temperature plasma modeling and simulation. *J. Micro/Nanopatterning, Mater., Metrol.* **22**, 041504–041504 (2023).
26. Mesbah, A. & Graves, D. B. Machine learning for modeling, diagnostics, and control of non-equilibrium plasmas. *J. Phys. D: Appl. Phys.* **52**, 30LT02 (2019).
27. Bonzanini, A. D., Shao, K., Graves, D. B., Hamaguchi, S. & Mesbah, A. Foundations of machine learning for low-temperature plasmas: methods and case studies. *Plasma Sources Sci. Technol.* **32**, 024003 (2023).
28. Bogaerts, A. et al. The 2020 plasma catalysis roadmap. *J. Phys. D: Appl. Phys.* **53**, 443001 (2020).
29. Laroussi, M. et al. Low-temperature plasma for biology, hygiene, and medicine: Perspective and roadmap. *IEEE Trans. Radiat. Plasma Med. Sci.* **6**, 127–157 (2021).
30. Kambara, M. et al. Science-based, data-driven developments in plasma processing for material synthesis and device-integration technologies. *Jpn. J. Appl. Phys.* **62**, SA0803 (2022).
31. Anirudh, R. et al. 2022 review of data-driven plasma science. *IEEE Trans. Plasma Sci.* **51**, 1750–1838 (2023).
32. Kapil, S. & Ali, M. Perspectives on artificial intelligence for plasma-assisted manufacturing in semiconductor industry. *IEEE Artificial Intelligence in Manufacturing Applications and Case Studies*, Chapter 4, 97–138 <https://doi.org/10.1016/B978-0-323-99135-3.00010-5> (2024).
33. Kawaguchi, S., Takahashi, K. & Satoh, K. Data-driven discovery of electron continuity equations in electron swarm map for determining electron transport coefficients in argon. *J. Phys. D: Appl. Phys.* **56**, 244003 (2023).
34. Zhong, L., Gu, Q. & Wu, B. Deep learning for thermal plasma simulation: Solving 1-d arc model as an example. *Comput. Phys. Commun.* **257**, 107496 (2020).
35. Mathews, A. et al. Uncovering turbulent plasma dynamics via deep learning from partial observations. *Phys. Rev. E* **104**, 025205 (2021).
36. Gergs, T., Borislovov, B. & Trieschmann, J. Efficient plasma-surface interaction surrogate model for sputtering processes based on autoencoder neural networks. *J. Vac. Sci. Technol. B* **40** (2022).
37. Kim, B. et al. Deep neural network-based reduced-order modeling of ion-surface interactions combined with molecular dynamics simulation. *J. Phys. D: Appl. Phys.* **56**, 384005 (2023).
38. Krüger, F., Gergs, T. & Trieschmann, J. Machine learning plasma-surface interface for coupling sputtering and gas-phase transport simulations. *Plasma Sources Sci. Technol.* **28**, 035002 (2019).
39. Shen, Y., Song, Z. & Kusiak, A. Enhancing the generalizability of predictive models with synergy of data and physics. *Meas. Sci. Technol.* **33**, 034002 (2021).
40. Guo, F. et al. Improving the out-of-sample generalization ability of data-driven chiller performance models using physics-guided neural network. *Appl. Energy* **354**, 122190 (2024).
41. Bikmukhametov, T. & Jäschke, J. Combining machine learning and process engineering physics towards enhanced accuracy and explainability of data-driven models. *Comput. Chem. Eng.* **138**, 106834 (2020).
42. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Computational Phys.* **378**, 686–707 (2019).
43. Rackauckas, C. et al. Universal differential equations for scientific machine learning. Preprint at <https://arxiv.org/abs/2001.04385> (2020).
44. Shen, C. et al. Differentiable modelling to unify machine learning and physical models for geosciences. *Nat. Rev. Earth Environ.* **4**, 552–567 (2023).
45. ElGazzar, A. & van Gerven, M. Universal differential equations as a common modeling language for neuroscience. <https://arxiv.org/abs/2403.14510> (2024).
46. Silvestri, M., Baldo, F., Misino, E. & Lombardi, M. An analysis of universal differential equations for data-driven discovery of ordinary differential equations. In: *Computational Science—ICCS 2023* (eds Mikyška, J. et al.) 353–366 (Springer Nature Switzerland, 2023).
47. Lima, F. A. R. et al. Improved modeling of crystallization processes by universal differential equations. *Chem. Eng. Res. Des.* **200**, 538–549 (2023).
48. Kuwahara, B. & Bauch, C. T. Predicting COVID-19 pandemic waves with biologically and behaviorally informed universal differential equations. *Heliyon* **10** (2024).
49. Smadbeck, P. & Kaznessis, Y. N. A closure scheme for chemical master equations. *Proc. Natl Acad. Sci.* **110**, 14261–14265 (2013).
50. MacNamara, S., Burrage, K. & Sidje, R. B. Multiscale modeling of chemical kinetics via the master equation. *Multiscale Modeling Simul.* **6**, 1146–1168 (2008).
51. Langmuir, I. The adsorption of gases on plane surfaces of glass, mica and platinum. *J. Am. Chem. Soc.* **40**, 1361–1401 (1918).
52. Shi, C., Yang, X., Zhang, J. & Zhou, T. Stochastic modeling of the mrna life process: A generalized master equation. *Biophys. J.* **122**, 4023–4041 (2023).
53. Pagnini, G., Mura, A. & Mainardi, F. Generalized fractional master equation for self-similar stochastic processes modelling anomalous diffusion. *Int. J. Stoch. Anal.* **2012**, 427383 (2012).
54. Donnelly, V. M. Reactions of fluorine atoms with silicon, revisited, again. *J. Vac. Sci. Technol. A* **35** (2017).
55. Athavale, S. D. & Economou, D. J. Molecular dynamics simulation of atomic layer etching of silicon. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **13**, 966–971 (1995).
56. Tan, S. et al. Highly selective directional atomic layer etching of silicon. *ECS J. Solid State Sci. Technol.* **4**, N5010 (2015).
57. Chen, R. T., Rubanova, Y., Bettencourt, J. & Duvenaud, D. K. Neural ordinary differential equations. *Adv. Neural Inform. Process. Syst.* **31** (2018).
58. Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* Vol. 1 (Elsevier, 1992).
59. Weidlich, W. & Haag, G. *Concepts and Models of A Quantitative Sociology: the Dynamics of Interacting Populations* Vol. 14 (Springer Science & Business Media, 2012).
60. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **4**, 251–257 (1991).
61. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control, Signals Syst.* **2**, 303–314 (1989).
62. Vella, J. R., Humbird, D. & Graves, D. B. Molecular dynamics study of silicon atomic layer etching by chlorine gas and argon ions. *J. Vac. Sci. Technol. B: Nanotechnol. Microelectron.: Mater. Process. Meas. Phenom.* **40**, 023205 (2022).
63. Tsaur, B., Matteson, S., Chapman, G., Liao, Z. & Nicolet, M.-A. Depth dependence of atomic mixing by ion beams. *Appl. Phys. Lett.* **35**, 825–828 (1979).
64. Haff, P. & Switkowski, Z. Ion-beam-induced atomic mixing. *J. Appl. Phys.* **48**, 3383–3386 (1977).
65. Andersen, H. The depth resolution of sputter profiling. *Appl. Phys.* **18**, 131–140 (1979).
66. Ho, P. Effects of enhanced diffusion on preferred sputtering of homogeneous alloy surfaces. *Surf. Sci.* **72**, 253–263 (1978).
67. Collins, R. & Carter, G. A first order diffusion approximation to atomic redistribution during ion bombardment of solids: I. Infinite range approximation. *Radiat. Eff.* **54**, 235–242 (1981).
68. Carter, G., Collins, R. & Thompson, D. A first order diffusion approximation to atomic redistribution during ion bombardment of solids: II. Finite range approximation. *Radiat. Eff.* **55**, 99–110 (1981).
69. Ohrlein, G. S. et al. Surface science issues in plasma etching. *IBM J. Res. Dev.* **43**, 181–197 (1999).
70. Kanarik, K. J. et al. Overview of atomic layer etching in the semiconductor industry. *J. Vac. Sci. Technol. A* **33** (2015).

71. Park, S.-D., Min, K.-S., Yoon, B.-Y., Lee, D.-H. & Yeom, G.-Y. Precise depth control of silicon etching using chlorine atomic layer etching. *Jpn. J. Appl. Phys.* **44**, 389 (2005).
72. Humbird, D. & Graves, D. B. Improved interatomic potentials for silicon–fluorine and silicon–chlorine. *J. Chem. Phys.* **120**, 2405–2412 (2004).
73. Sherpa, S. D., Ventzek, P. L., Lee, M., Hwang, G. S. & Ranjan, A. New insight into desorption step by Ar⁺ ion-bombardment during the atomic layer etching of silicon. *J. Vac. Sci. Technol. A* **36** (2018).
74. Coburn, J. W. & Winters, H. F. Ion- and electron-assisted gas-surface chemistry—an important effect in plasma etching. *J. Appl. Phys.* **50**, 3189–3196 (1979).
75. Sycheva, A. A., Voronina, E. N., Rakhimova, T. V. & Rakhimov, A. T. Argon clustering in silicon under low-energy irradiation: molecular dynamics simulation with different ar–si potentials. *J. Vac. Sci. Technol. A* **36** (2018).
76. Kim, D. H. et al. Molecular dynamics simulation of silicon dioxide etching by hydrogen fluoride using the reactive force field. *ACS Omega* **6**, 16009–16015 (2021).
77. Ohta, H. & Hamaguchi, S. Molecular dynamics simulation of silicon and silicon dioxide etching by energetic halogen beams. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **19**, 2373–2381 (2001).
78. Hanson, D. E., Voter, A. & Kress, J. Molecular dynamics simulation of reactive ion etching of Si by energetic Cl ions. *J. Appl. Phys.* **82**, 3552–3559 (1997).
79. Shim, S., Vella, J. R., Draney, J. S., Na, D. & Graves, D. B. An examination of the performance of molecular dynamics force fields: silicon and silicon dioxide reactive ion etching. *J. Vac. Sci. Technol. A* **42** (2024).
80. Zhang, R., Zhao, Y. & Teo, B. K. Fluorination-induced back-bond weakening and hydrogen passivation on hf-etched si surfaces. *Phys. Rev. B* **69**, 125319 (2004).
81. Haag, G. Modelling with the Master Equation: Solution Methods and Applications in Social and Natural Sciences 1 edn (Springer Cham, 2017).
82. Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **48**, 571 (1976).
83. Ouyang, W., Saven, J. G. & Subotnik, J. E. A surface hopping view of electrochemistry: non-equilibrium electronic transport through an ionic solution with a classical master equation. *J. Phys. Chem. C* **119**, 20833–20844 (2015).
84. Hao, Q., Kim, P., Nam, S. K., Kang, S.-Y. & Donnelly, V. M. Real-time monitoring of atomic layer etching in Cl₂/Ar pulsed gas, pulsed power plasmas by optical emission spectroscopy. *J. Vac. Sci. Technol. A* **41** (2023).
85. Sankaran, A. & Kushner, M. J. Integrated feature scale modeling of plasma processing of porous and solid SiO₂. I Fluorocarbon etching. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **22**, 1242–1259 (2004).
86. Palov, A. P., Balint-Kurti, G. G., Voronina, E. N. & Rakhimova, T. V. Sputtering of Si by Ar: a binary collision approach based on quantum-mechanical cross sections. *J. Vac. Sci. Technol. A* **36** (2018).
87. Barsukov, Y. V. et al. Enhanced silicon nitride etching in the presence of f atoms: Quantum chemistry simulation. *J. Vac. Sci. Technol. A* **36** (2018).
88. Dwivedi, O. D., Barsukov, Y., Jubin, S., Vella, J. R. & Kaganovich, I. Orientation-dependent etching of silicon by fluorine molecules: a quantum chemistry computational study. *J. Vac. Sci. Technol.* **41** (2023).
89. Graves, D. B. & Humbird, D. Surface chemistry associated with plasma etching processes. *Appl. Surf. Sci.* **192**, 72–87 (2002).
90. Alves, L., Bogaerts, A., Guerra, V. & Turner, M. Foundations of modelling of nonequilibrium low-temperature plasmas. *Plasma Sources Sci. Technol.* **27**, 023002 (2018).
91. Repin, D. & Petrov, T. Automated deep abstractions for stochastic chemical reaction networks. *Inf. Comput.* **281**, 104788 (2021).
92. Bortolussi, L. & Palmieri, L. in *Computational Methods in Systems Biology* (eds Češka, M. & Šafránek D.) 21–38 (Springer International Publishing, 2018).
93. Bortolussi, L. & Cairoli, F. in *Quantitative Evaluation of Systems* (eds Parker, D. & Wolf, V.) 259–276 (Springer International Publishing, 2019).
94. Lill, T. *Atomic Layer Processing: Semiconductor Dry Etching Technology* (John Wiley & Sons, 2021).
95. Balakrishnan, V. On a simple derivation of master equations for diffusion processes driven by white noise and dichotomic markov noise. *Pramana* **40**, 259–265 (1993).
96. Gasvoda, R. J., Zhang, Z., Wang, S., Hudson, E. A. & Agarwal, S. Etch selectivity during plasma-assisted etching of sio₂ and sinx: transitioning from reactive ion etching to atomic layer etching. *J. Vac. Sci. Technol. A* **38** (2020).
97. Gasvoda, R. J., van de Steeg, A. W., Bhowmick, R., Hudson, E. A. & Agarwal, S. Surface phenomena during plasma-assisted atomic layer etching of sio₂. *ACS Appl. Mater. Interfaces* **9**, 31067–31075 (2017).
98. Ho, P., Lewis, J., Wildman, H. & Howard, J. K. Auger study of preferred sputtering on binary alloy surfaces. *Surf. Sci.* **57**, 393–405 (1976).
99. Einstein, A. et al. On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat. *Ann. Phys.* **17**, 208 (1905).

Acknowledgements

This material is based upon work supported in part by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences under Award No. DE-SC0024472 and Award No. DE-SC0024474.

Author contributions

S.N. and A.M. conceptualized the work. S.N. ran the computational experiments. S.N. and A.M. wrote the main manuscript. All authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-025-01677-4>.

Correspondence and requests for materials should be addressed to Ali Mesbah.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025