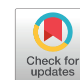
Contents lists available at ScienceDirect

Chemical Engineering Research and Design

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



Microscopic modeling and optimal operation of plasma enhanced atomic layer deposition

Yangyao Ding[a], Yichi Zhang[a], Gerassimos Orkoulas[c], Panagiotis D. Christofides[a,b,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA
[b] Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA
[c] Department of Chemical Engineering, Widener University, Chester, PA 19013, USA

ARTICLE INFO

Article history:
Received 28 April 2020
Received in revised form 6 May 2020
Accepted 12 May 2020
Available online 23 May 2020

Keywords:
Plasma enhanced atomic layer deposition
Microscopic modeling
Kinetic Monte Carlo modeling
Density functional theory
Neural networks

ABSTRACT

Plasma enhanced atomic layer deposition (PEALD) is one of the most widely adopted deposition methods used in the semiconductor industry. It is chosen largely due to its superior ability to produce ultra-thin high-k dielectric films, which are needed for the further miniaturization of microelectronic devices with the pace of Moore's Law. In contrast to the traditional thermal atomic layer deposition (ALD) method, PEALD allows for high deposition growth per cycle (GPC) under low operating temperature with the help of high energy plasma species. Despite the experimental effort in finding new precursors and plasmas, the detailed surface structures and reaction mechanisms in various PEALD processes remain hard to understand because of the limitation of current in-situ monitoring techniques and the deficiency of the first-principles based analysis. Therefore, in this work, an accurate, yet efficient kinetic Monte Carlo (kMC) model and an associated machine learning (ML) analysis are proposed to capture the surface deposition mechanism and to propose optimal operating conditions of the HfO_2 thin-film PEALD using tetrakis-dimethylamino-Hafnium (TDMAHf) and oxygen plasma. Density Functional Theory (DFT) calculations are performed to obtain the key kinetic parameters and the structural details, subsequently employed in the kMC model. After the kMC model is validated by experimental data, a database is generated to explore a variety of precursor partial pressure and substrate temperature combinations using the kMC algorithm. A feed-forward Bayesian regularized artificial neural network (BRANN) is then constructed to characterize the input–output relationship and to investigate the optimal operating conditions.

© 2020 Institution of Chemical Engineers. Published by Elsevier B.V. All rights reserved.

1. Introduction

In recent decades, atomic layer deposition (ALD) has been widely utilized as a modern deposition technique in the semiconductor manufacturing industry. It has been developed to meet the demand of producing conformal thin-films in high aspect ratio structures for the metal-oxide-semiconductor field-effect transistors (MOSFETs) in the NAND-type flash memory devices (Schuegraf et al., 2013). To allow for the device size reduction to follow the Moore's Law and to enhance the performance of the microelectronic components, the thickness and width of the traditional oxide films (e.g., SiO_2, Si_3O_4 and Al_2O_3) have been constantly reduced to decrease the overall size of the transistor and to increase the gate capacitance effectiveness (Nalwa, 2002). Nevertheless, as the thickness of the oxides scales below 20 Å, there exists a significant increase of leakage current through the oxide film due to quantum tunneling, where a substantial portion of the electrons passes through the closed transistor gate caused by the low potential barrier of the extremely thin oxide films (Jeong et al., 2016). Thus, the spiking induced power consumption and the lack of device reliability limit further part miniaturization. As a

* Corresponding author.
E-mail address: pdc@seas.ucla.edu (P.D. Christofides).
https://doi.org/10.1016/j.cherd.2020.05.014
0263-8762/© 2020 Institution of Chemical Engineers. Published by Elsevier B.V. All rights reserved.

result, replacing the traditional oxide film with high dielectric constant (high-κ) materials, where large enough potential barriers can be achieved with a similar thickness scale, has been extensively investigated (Ribes et al., 2005). Nevertheless, the manufacturing of those high-κ thin-films using thermal ALD has faced challenges of low-throughput and high operating cost (Joo and Rossnagel, 2009). The primary reason for these problems is the extremely slow reaction kinetics at low temperatures because of the high activation energy barrier of the traditional precursors, thus often requiring high temperature for energy input. To overcome this problem, plasma is introduced as an enhancement to the originally neutral precursors, in which free radicals and ions are created so that the deposition reactions can go through lower energy barrier pathways (Won et al., 2005).

Plasma enhanced atomic layer deposition (PEALD) is designed as a variation of ALD, which utilizes high-energy plasma species to replace the extremely high operating temperature. In a PEALD process, the substrate surface is alternatively exposed to heavy-metal precursor pulses and plasma-enhanced precursor pulses, with purging pulses with inert gas in between. Each of the precursor pulses is called a half-cycle, in which the associated precursors react with the substrate surface in a self-limiting manner. Given sufficient precursor dosage, appropriate half-cycle time, and operating conditions, each half-cycle ideally results in a conformal surface with full coverage. While in the purge cycle, an inert gas is introduced to remove the left-over precursor gases and by-products from the last half-cycle to avoid undesirable multi-layer deposition (George, 2009). As a result, the PEALD method can ideally grow uniform high-κ dielectric thin-films layer-by-layer in a controlled but rapid manner, while avoiding high operating temperature and reducing energy consumption (Ishikawa et al., 2017).

In the past few years, there has been a growing number of PEALD works on both laboratory and industrial scales to investigate new types of thin-film materials, novel precursors, and their respective reaction mechanisms which enable improved PEALD operation. High-k dielectric thin-film materials like HfO_2 (Liu et al., 2005), TiO_2 (Kukli et al., 2000) and ZrO_2 (Yun et al., 2004) are typically investigated due to their good band alignment to silicon gate, the high thermal stability, and the capability of offering high mobility for charge carriers. A variety of precursor species have been investigated for both the neutral heavy particle pulse and the plasma pulse for the deposition of the aforementioned high-k thin films. Taking HfO_2 as an example, for the first half-cycle, metal organic precursors are extensively used because of the low activation energy from the H-N hydrogen bonds. There has been a variety of research works for popular Hf-based metal organic precursors including tetrakis(ethylmethylamido) hafnium (TEMAH) (Kukli et al., 2000), tetrakis(dimethylamido) hafnium (TDMAH) (Liu et al., 2003), tetrakis(diethylamido) hafnium (TDEAH) (Shi et al., 2011), and others. For the plasma half-cycle, common oxide precursors involve H_2O plasma (Kanomata et al., 2014), O_2 (Rai et al., 2010) plasma, and O_3 plasma (Rai et al., 2010). Despite the various advantages associated with the PEALD process, the experimental studies still face obstacles in operating costs to explore a substantial amount of experimental conditions and in acquiring clear first-principles based understandings. In the reaction mechanisms domain, the surface precursor and plasma reactions are crucial to the deposition profile, but the detailed mechanisms remains partially unknown (Lee et al., 2006). In the process monitoring domain, the real-time measurements of the microscopic crystal structure and uniformity are limited. Currently, in-depth micro-structure analysis methods rely on the scanning electron microscope (SEM) and the scanning tunneling microscope (STM), which are destructive to the deposited film (Schwille et al., 2017). Also, the in-situ monitoring technologies like the spectroscopic ellipsometry are feasible but are not robust because of the complex optical interaction (Pittal et al., 1993). Therefore, the development of an accurate microscopic surface deposition model is critical to help understand the PEALD process and is the first fundamental step to construct an all-inclusive model to characterize the entire PEALD process.

There has been a variety of attempts to describe the deposition reactions on the substrate surface beyond the Langmuir adsorption model (Travis and Adomaitis, 2013, 2014). Molecular Dynamics (MD) has traditionally been utilized where an ab initio model tracks the movement of all particles. Nevertheless, such simulations are too computationally expensive for an industrial-scale system (Battaile and Srolovitz, 2002). Thus, the application of the kinetic Monte Carlo (kMC) method, specifically the reject-free kMC method, which follows a single event at a time, dramatically boosts the computational efficiency while preserving the model validity (Rey et al., 1991; Elliott and Greer, 2004; Dkhissi et al., 2008). Many groups have incorporated kMC to simulate a variety of surface deposition processes. Knoops et al. (2010) adopted the empirical recombination probabilities for a standard kMC algorithm to simulate a PEALD process. Shirazi and Elliott (2014) modeled a HfO_2 ALD on a small-scale surface using a thorough set of first-principles-based reactions for the kMC model. Crose et al. (2018) recently proposed a multiscale CFD modeling approach, which integrated a microscopic kMC model and a macroscopic CFD model for plasma enhanced chamber vapor deposition (PECVD). Ding et al. (2019a) formulated a 3D lattice model that represented the competitive reaction pathway of bis(tertiary-butylamino)silane (BTBAS) precursor for the thermal SiO_2 thin-film ALD. Despite the existing efforts to describe the microscopic domains, there has been no kMC model constructed for PEALD of HfO_2 thin-film that is efficient enough to be applied for an industrial-scale PEALD process.

Still, there are some disadvantages associated with the general kMC models. In particular, the kMC model may be computationally expensive to be implemented in the real-time manufacturing edge analytics. Also, the algorithm does not have a closed-form solution, which makes it hard to be incorporated in the model-based optimization and control. In order to deal with those problems, computationally efficient data-driven approaches can be utilized for the deposition processes. For example, Ding et al. (2019b) used a linear parameter varying model to characterize and optimize a thermal ALD reactor. Nevertheless, traditional statistical and machine learning methods like the ordinary least square regression (OLSR) often fail to reach good enough accuracy because of the complex non-linear input–output relationship in chemical engineering processes. As a result, with the enlightenment of the neural structures and the increasing computational power in recent decades, neural networks (NN) have been extensively researched in the domain of machine learning. A neural network utilizes an interconnected group of artificial neurons, each handling a simple mathematical relationship for the overall complex information processing based on a connectionist approach. Depending on the external and internal information that flows through the network, the NN structure can be adapted to perform classification or regression

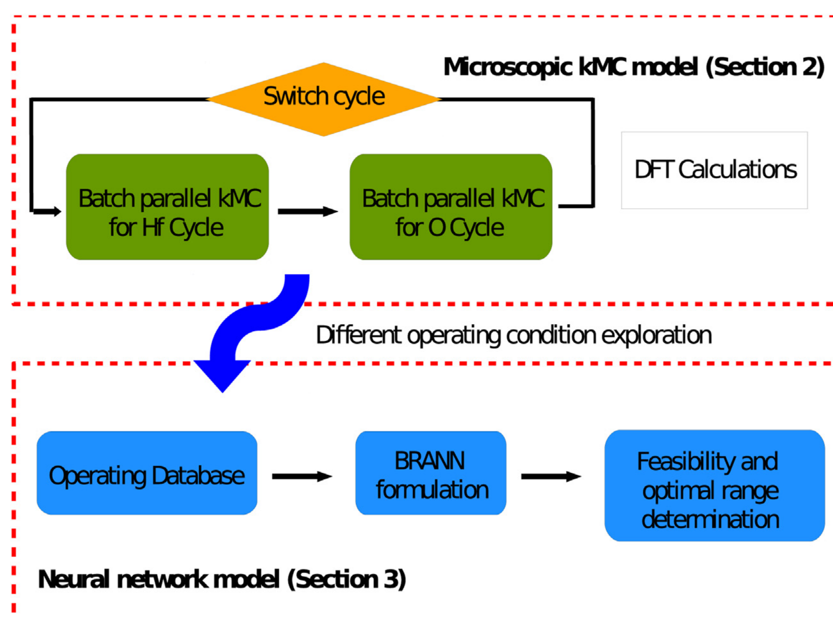


Fig. 1 – Workflow demonstrated in this paper and the respective section of each component.

on complicated information processing. Zhang et al. (2020) adopted the NN-assisted model to apply PI and batch-to-batch control in order to deal with the disturbances in the furnace ALD manufacturing. Rasoulian and Ricardez-Sandoval (2015) and Kimaev and Ricardez-Sandoval (2019) adopted various advanced neural network algorithms for the thin-film deposition processes and incorporated corresponding model-based controls. These previous works have demonstrated the enormous potential of the possible development of sophisticated online optimizations and advanced control schemes for deposition processes (Armaou and Christofides, 1999).

Therefore, in this paper, a computationally efficient, yet sufficiently rigorous microscopic kMC model has been proposed for the PEALD process with TDMAH and O_2 plasma as precursors. Specifically, important structural details and reaction parameters are computed from the density functional theory (DFT) calculation. Depending on the calculated parameters, we adopt a 3D kMC algorithm to capture the surface reaction mechanism and the structure detail of the HfO_2 thin-film. From the developed simulation, a meaningful range of experimental conditions is explored to construct a deposition profile database. Based on this dataset, neural network regression is utilized to map the inlet parameter combinations to the film structures and growth profiles, and to figure out the optimal operating conditions according to throughput and economic concern. The workflow of this paper is shown in Fig. 1.

2. PEALD microscopic modeling

This section addresses the microscopic film growth model of the PEALD. First, the HfO_2 structure detail and corresponding kMC lattice representation are discussed. Next, a selected set of reaction mechanisms that are critical to surface deposition, and their respective rate calculations are explained. After that, the kMC algorithm time progression and event selection are formulated in detail. In addition, the DFT calculations are introduced for the determination of the important process and structural parameters. Finally, the automated workflow with cloud distributed computing is demonstrated.

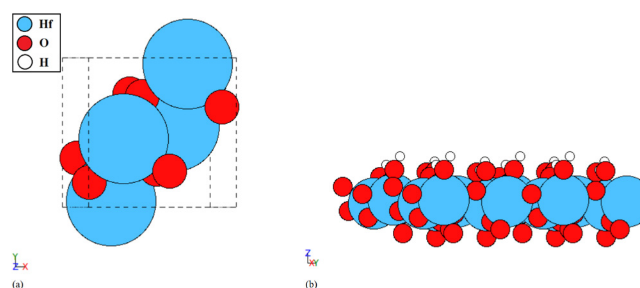


Fig. 2 – (a) Optimized monoclinic HfO_2 unit cell in the bulk, and (b) relaxed OH-terminated HfO_2 2 × 2 ×1 slab.

2.1. *Structural characterization of the deposited HfO_2*

There are various crystalline structures that HfO_2 can take on, including tetragonal ($P4_2/nmc$), cubic ($Fm\bar{3}m$), monoclinic ($P2_1/c$), orthorhombic (Pbca and Pnma), etc. (Zeng et al., 2014). Although it is known that the PEALD method often leads to an amorphous structure that is close to some crystalline phase, the exact phase is critically dependent on the deposition temperature. At very high temperatures, the HfO_2 often forms in the cubic phase, whereas the surface is often in the monoclinic phase at low temperatures (Kumar et al., 2017). In this paper, since we are investigating the PEALD at low temperature, we consider a surface structure that takes the monoclinic-alike structure. The (111) surface, known as one of the most common surfaces of Hafnia, is chosen as the deposition surface, and the growth direction is shown as the z-direction in Fig. 2 (a). Also, the hydroxyl-terminated surface, which is commonly prepared by the surface treatment with piranha solution and then dilute HF solution to remove native oxide and organic compounds, is chosen as the initial substrate surface, which is shown in Fig. 2 (b). Although it seems to be the most realistic if we adopt the exact monoclinic structure angle and bond length, the complex reaction pathways and structural detail in this real lattice make the kMC formulation computationally and conceptually challenging. As a result, we adopt a 3D triangular on-lattice model, developed in Ding et al. (2019a), as an approximation of the local tetrahedron alike crystal structure. In the approximated lattice structure, the atoms are consid-

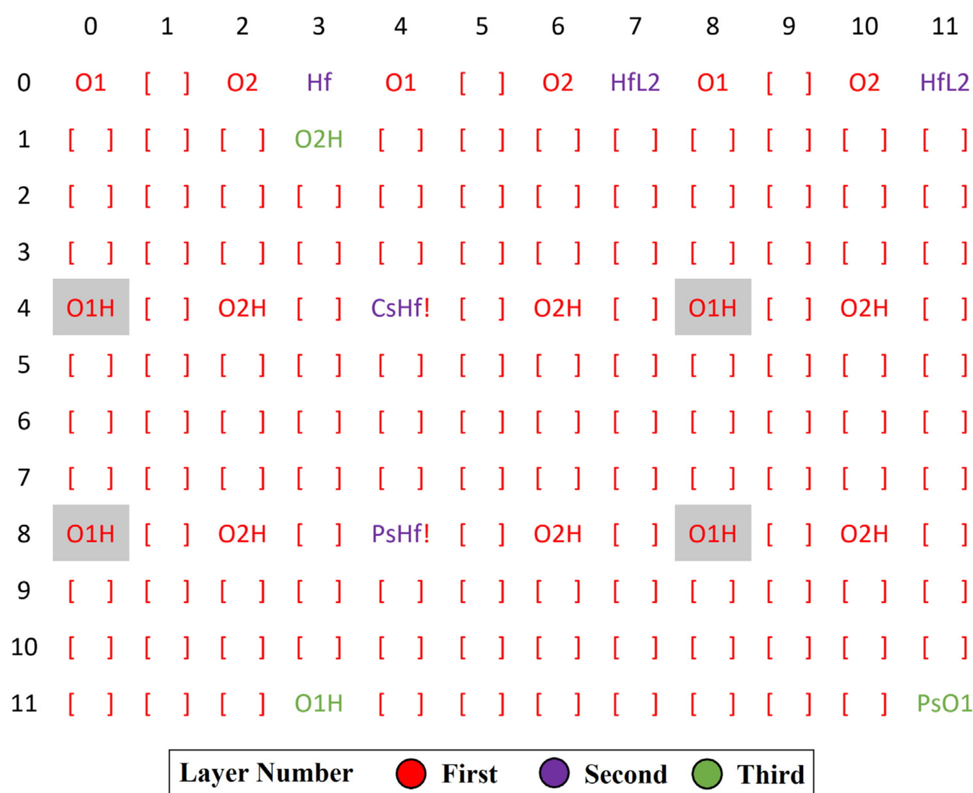


Fig. 3 – Top view of a 3-layer 12 × 12-site miniature demonstration of the full lattice used in the model. The three layers and the species on the lattice are shown using different colors and symbols, respectively. Grey background denotes the sterically hindered sites. The exclamation mark (!) denotes the OH group directly underneath.

ered on a crystal lattice, where the bonds connecting the Hf and the adjacent O atom all have 90 degree angles with respect to each other. This lattice structure assumption successfully preserves the key structural characteristics. The exemplary lattice miniature of our simulation is shown in Fig. 3, where HfL2 is the deposited Hf with two unreacted methylamino ligands, and PsHf and CsHf are physisorbed and chemisorbed precursors, respectively. PsO1 or PsO2 are the physisorbed oxygen species which remain to be oxidized. This top view is shown to closely resembles the surface of the real PEALD lattice, as shown in Fig. 2(b). Also, the assumption leads to a similar cycle repetition as found in the realistic structure. The structural validation will be further demonstrated in the model validation section. The actual lattice used in the simulation has 1200 × 1200 sites per layer with a height that equals to the number of cycles simulated, which is proven to be large enough to be size-independent (Huang et al., 2010).

2.2. *Reaction mechanism*

As introduced in the general PEALD process, a full cycle is comprised of two half-cycles with purging cycles in between. In each half-cycle and the beginning of its purging step, a specific or a set of specific precursors will contact and react with the substrate surface, which deposits the desired element onto the film. Tetrakis-dimethylamino-Hafnium (TDMAHf) is considered to be the precursor for the Hafnium half cycle, and molecular oxygen ($^3O_2$), singlet molecular oxygen ($^1O_2$), and triplet atomic oxygen (^3O), generated from the plasma chamber, are considered to be the precursors of the oxygen half-cycle. The singlet atomic oxygen (^1O), according to the Tinck and Bogaerts (2011), recombines at a high rate in the gas transport in the main reactor, and will thus not be considered for the surface reaction. The detailed reaction mechanisms of both cycles were reported by Shirazi and Elliott (2014) and Jeon and Won (2008). However, the implementation of the full reaction set, which follows all the possible mechanisms of ligand rotation, proton diffusion, and other processes, will dramatically increase the computational demand, and thereby make the simulation of an industrial-sized wafer infeasible. Therefore, key reactions are selected with essential details on the proton transfer and structural influence considered, which would still constitute a full deposition reaction set and reproduce the experimentally reported growth rates and potential defects. The selected reaction mechanisms are explained below.

The first half-cycle, also referred to as the Hf-Cycle, contains physisorption, desorption, and a two-step dissociative chemisorption, which is shown in the top part of Fig. 4. The precursor TDMAHf is exposed to the hydroxyl group terminated Hf surface, as shown in Fig. 4(a). According to Han et al. (2011), the two oxygen atoms in a cell of a surface oxide typically have different electronegativities, which is caused by the distinctive extent of exposure from the local bonding geometry. This conclusion is also implemented for HfO_2 in this work, in which the more electronegative oxygen atom is more likely to be attacked by the precursor. Therefore, as indicated by Fig. 4(1), the nitrogen atom in the precursor particle enables strong physisorption onto the more electronegative hydroxyl group through an H-bond. After the physisorption, the precursor undergoes the first chemisorption step, as indicated by Fig. 4(2), where the surface H from the attacked O will transfer across the ligands on the precursor, as shown in Fig. 4(a). Eventually, the Hf is chemically bonded to the surface, releasing a dimethylamine (DMA) group, shown in Fig. 4(c). After this first dissociative chemisorption step, more protons may

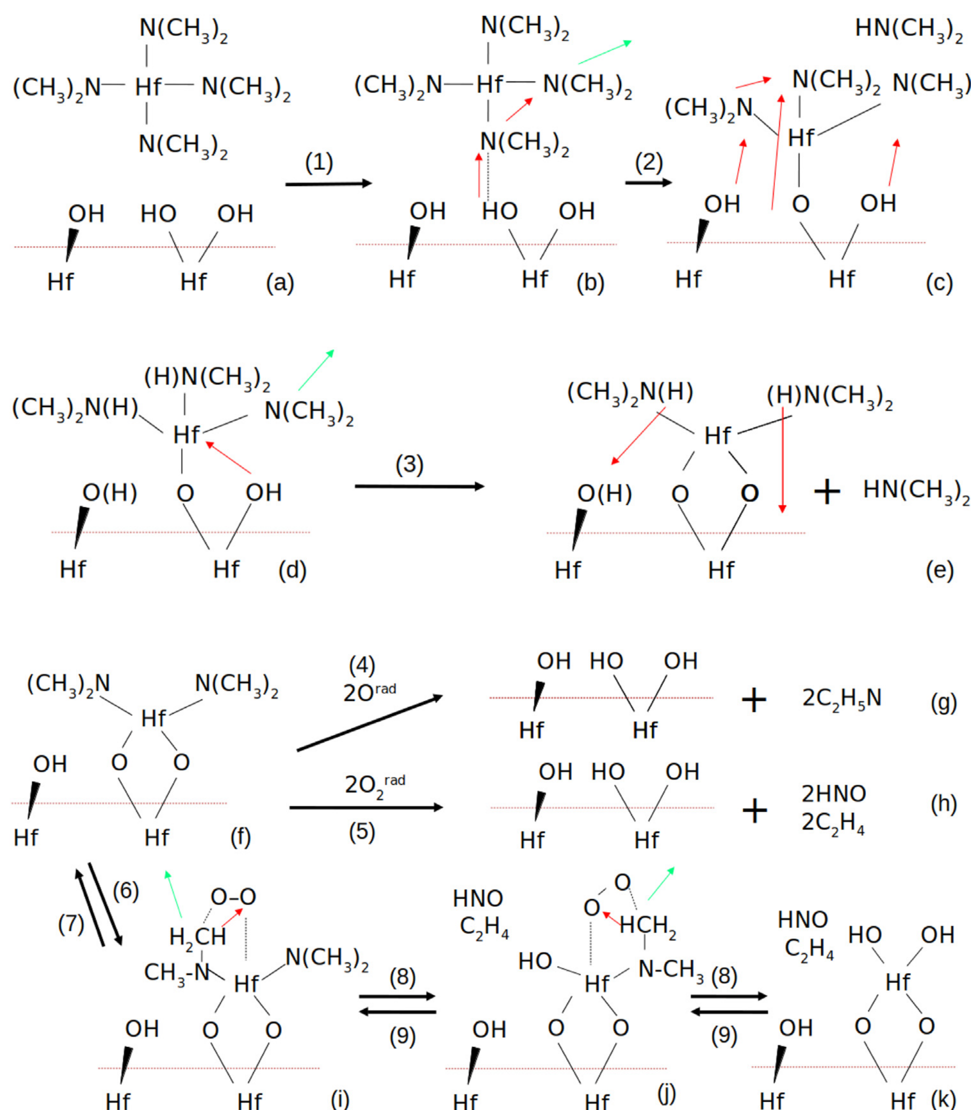


Fig. 4 – Illustration of the selected representative reaction set. The top part is the Hf-Cycle and the bottom part is the O-Cycle. The black arrows represent the overall reaction pathway, the red arrows represent the possible proton diffusion, and the (*H*)s represent the potentially protonated positions. The number labels refer to the events and the alphabetic labels refer to the states. Please note that the bond lengths are not exact and drawn for descriptive purposes.

be transferred from the nearby substrate surface toward the precursor, making some of the ligands in the protonated state. Subsequently, one of the remaining DMA groups reacts with the adjacent hydroxyl group from the neighbour Hf atom, indicated by Fig. 4(3), where again the surface H from the attacked O will transfer across the ligands on the precursor, shown in Fig. 4(d), and release another DMA group, thus anchoring the Hf on the surface. This reaction pathway retains the original surface orientation and results in a thermodynamically favorable structure. After the reaction, the remaining two DMA groups become the new surface species, and some ligand protons will return to the nearby substrate surface, shown in Fig. 4(e).

The second half-cycle is the O-Cycle, or the plasma cycle. Due to the high energy of radicals, i.e., plasma species, they often undergo and encounter different reaction routes and energy barriers than the neutral non-plasma species, thereby facilitating the deposition process. In particular, the O-cycle contains the following processes: physisorption, abstraction and surface oxidation, and the respective reverse reaction. While the majority of the works ignore the deposition effect of the neutral oxygen gas (Melo et al., 2016; Heil et al., 2007), Provine et al. (2016) stresses the importance of the oxygen gas contribution in the oxygen plasma. As a result, we take all of the potentially important species from the plasma reaction into consideration in our model: the molecular oxygen ($^3O_2$), the singlet molecular oxygen ($^1O_2$), and the triplet atomic oxygen (^3O). According to the kinetics reported by Jeon and Won (2008), the oxidation of the surface DMA group by the radicals, $^1O_2$ and ^3O, are extremely fast and kinetically barrierless. Therefore, as soon as those species are physisorbed onto the surface, they instantaneously oxidize and replace the surface DMA group with hydroxyl groups, as shown in Fig. 4(4) and (5), leaving the by-product gas of ethanimine compounds or the nitroxyl and ethane, respectively, indicated by Fig. 4(g) and (h). For the non-plasma species, $^3O_2$, however, it is essential to understand the competition of the surface chemical reactions because of the high transition state energy barrier. First, the $^3O_2$ physisorbs onto the DMA terminated surface, as indicated by the step Fig. 4(6), to form a ring-shaped physisorbed complex, as shown in Figure 4 (i). This physisorbed complex could be abstracted from the surface, as shown in Fig. 4(7), because the addition of molecular oxygen may increase the stability of the DMA groups attached to the surface (Epling et al., 1998).

Moving forward, the physisorbed complex may go through a series of oxidation steps, shown in Fig. 4(8), producing nitroxyl and ethane in each step indicated by Fig. 4(j) and (k). The reverse reaction, indicated by Fig. 4(9), is associated with high activation energy. To characterize the aforementioned O-Cycle reaction mechanisms, we combine our own DFT calculation with some reported reaction intermediate complex and transition state profile reported by Jeon and Won (2008).

It is important to note that the proton transfer back and forth from the ligand and the surface discussed in this section has a very minimal energy barrier and thus will happen extremely fast. Also, those events depend on the complex local coordination number and ligand rotation state, and thus, they are impossible to be deterministically captured. As a result, the possible protonated states will be randomly captured in the kMC, which will slightly influence the reaction energy. In addition, steric hindrance is an important concern that influences the surface site occupation. According to Shirazi and Elliott (2014), due to the bulkiness of the DMA groups, once a TDMAHf is physisorbed on the surface, indicated by Fig.. 4 (1), it will cause steric hindrance to the nearby sites, which means the closest hydroxyl groups will not be able to accept new precursor species, as roughly seen in Fig. 4(b). This blocking effect will continue to affect the local deposition environment until the surface DMA groups are oxidized, as shown in Fig. 4(g), (h), and (k). Those two specific considerations extract the important characteristic of the comprehensive model developed by Shirazi and Elliott (2014) and other experimental findings, which will help our model more closely reproduce the realistic PEALD process conditions while maintaining a computationally acceptable algorithm.

2.3. *Relative rate determination*

The reaction kinetics of the pathways discussed in the previous section is crucial to the kMC event selection scheme, which influences the competitions between the aforementioned reactions and the overall deposition profile. There are majorly two kinds of reactions in the adsorption process, the physisorption and the surface chemical reactions, that can be characterized by different governing theories.

The chemisorptions, which in our case include the dissociative chemisorption, the desorption, and the forward and reverse oxidation reactions are thermodynamically activated reactions, which can be described by the general Arrenhius-type equation governed by the transition state theory (TST) (Cortright and Dumesic, 2001). The reaction rate is described by the thermodynamic properties of the transition state complexes at the quasi-equilibrium. As a result, the Arrenhius-type reaction rate of the aforementioned reactions can be formulated as follows:

$$r_{rxn,i} = A_i \exp(\frac{-E_{a,i}}{k_b T}) \tag{1}$$

where i refers to the index of each individual reaction, $E_{a,i}$ is the activation energy of the transition state complex, and A_i is the pre-exponential factor, which is governed by the following relationship:

$$A_i = f_i^{TST}(\frac{k_b T}{h}) \tag{2}$$

where f_i^{TST} is the vibrational partition function ratio between the transition state complex and the reactants, k_b is the Boltzmann constant, and T is the temperature. The f_i^{TST} is around 1 for reactions happened on the slab surface (Shirazi and Elliott, 2014).

In contrast, the physisorption r_{phs} can be described by the gas collision theory, which is reasonable for the gas-surface athermal barrierless processes:

$$r_{phs} = \frac{p_i}{RT}\sqrt{\frac{8RT}{\pi M_i}} s_{c,i} N_a \sigma \tag{3}$$

where for species i, p_i is the partial pressure of the precursor species, either TDMAHf or oxygen radicals, R is the gas constant, T is the absolute temperature, M_i is molar mass, $s_{c,i}$ is the sticking coefficient, σ is the unit cell surface area, and N_a is the Avogadro number. The sticking coefficient of TDMAHf changes exponentially with varying surface temperature, which is reported by Dorsmann and Kleijn (2007).

2.4. *Kinetic Monte Carlo algorithm*

In order to simulate the above reaction pathways and to realistically represent the reaction rates and structural evolution, first-principles based modeling method is adopted. In this paper, we choose to utilize the n-fold kinetic Monte Carlo (kMC) algorithm in the framework proposed by earlier works (Ding et al., 2019a). kMC is a computational random sampling algorithm that, in this case, simulates the time progression of the surface deposition based on random variables and reaction kinetics. In particular, we focus on two aspects: event selection and time progression.

For the event selection, a random number, $\gamma_1 \in (0, 1]$, is chosen from the uniform distribution. To utilize γ_1, we define a collective set of possible events, which includes all the physical and chemical reactions. The total reaction rate r_{total} can be calculated as the sum of the unweighted rates calculated from the individual events:

$$r_{total} = \sum_{i=1}^{N} r^i_{unweighted} \tag{4}$$

where $r^i_{unweighted}$ represents the unweighted rate of the ith event within an N-event set, each calculated from the equations discussed in the previous section. Adopting the aforementioned equations, the event selection details of the two half-cycles can be computed as follows. For the Hf half-cycle, the total rate is:

$$r_{Hf,total} = r_{Hf,phs} + r_{Hf,des} + r_{Hf_a,che1} + r_{Hf_b,che2} \tag{5}$$

where $r_{Hf,phs}$ is the rate of TDMAHf physisorption, $r_{Hf,des}$ is the rate of TDMAHf desorption, and $r_{Hf_a,che1}$ together with $r_{Hf_b,che2}$ are the dissociative chemisorption rates of the TDMAHf first and second chemisorption, respectively. In contrast, for the O-Cycle, the total rate is:

$$r_{O,total} = r_{3O,phs} + r_{1O_2,phs} + r_{3O_2,phs} + r_{3O_2,des} + r_{3O_2,oxidation} \tag{6}$$

where $r_{3O,phs}$, $r_{1O_2d,phs}$, and $r_{3O_2d,phs}$ are the physisorption rates of atomic oxygen radicals, oxygen radicals and ground-state oxygen gas, respectively. $r_{3O_2,des}$ and $r_{3O_2,oxidation}$ are the desorption and forward oxidation reaction, respectively. Chemisorption of the radical gas is not included in the event selection because the energies associated with atomic oxy-

gen and oxygen radicals are extremely high. Thus, the barrier from the reaction complex to transition state can be easily overcome and the reaction can be considered instantaneous.

The second aspect of the kMC algorithm is the computation of the time progression to model the transient behavior of the surface deposition, utilizing a random number $\gamma_1 \in (0, 1]$. The time of each selected event is governed by the total rate as follows:

$$\Delta t = \frac{-\ln \gamma_2}{r_{total}} \tag{7}$$

Starting from $t = 0$, the simulation time clock is consecutively advanced by Δt of each event until it reaches the predefined half-cycle time.

Nevertheless, in order to reproduce the realistic competitions between the surface physical and chemical reactions, we need to consider a variation from the standard kMC, where the surface species and vacant site concentrations are taken into consideration. Thus, the normalized event indicators under the competition of reaction pathways and directions are calculated by dividing the concentration-weighted reaction rates with respect to the total weighted rate as follows:

$$l_{Hf,i} = \frac{\sum_{j=1}^{i} r^{j}_{unweighted} R_j}{\sum_{k=1}^{N} r^{k}_{unweighted} R_k}, \quad i = 1, \ldots, N \tag{8}$$

where $l_{Hf,i} \in (0, 1]$ represents the normalized indicator of the ith event in the surface reaction event set, $r^{j}_{unweighted}$ and $r^{k}_{unweighted}$ are the unweighted chemical reaction rate for the jth and kth event, respectively, and N is the total number of events in the hf-cycle or the o-cycle reaction event set. R is the number of available reactants for each associated reaction. Physisorption considers the acceptable empty sites, forward chemisorption and desorption consider the physisorbed species, and reverse reactions consider the chemisorbed or oxidized species. A second random number γ_2 is chosen for event selection. If the number lies in the ith interval, between the normalized indicators l_{i-1} to l_i, the ith event is chosen for execution. It will be demonstrated in the model validation section that this tailored kMC algorithm takes in the first-principles-based characteristic as in the Langmuir adsorption model and the structural details to successfully reproduce the complex deposition process.

2.5. *DFT and thermodynamic calculations*

Parts of the first-principles analysis are reported for the interaction of the HfO_2 substrate surface with TDMAHf and O_2 plasma. However, to fully understand the exact surface configurations and deposition profile, we adopt the molecular visualization toolbox ASE and Density Functional Theory (DFT) calculation with the Gaussian09 software package to extract all necessary reaction mechanism, thermodynamic properties, and structural detail. Such information would help the microscopic kMC model to reproduce realistic film growth behavior. (Frisch et al., 2000; Liu and Liu, 2018; Zhang et al., 2019).

DFT is an approximation approach to the solution of the Schrödinger equation of the many-body system:

$$\hat{H}\Psi = E\Psi \tag{9}$$

where Ψ is the state vector of the quantum system, and E is the energy eigenvalue associated with the quantum state Ψ. In addition, $\hat{H}$ is the Hamiltonian operator, which consists of the nucleic kinetic energy $\hat{T}_{nuc}(R)$, the coulombic potential between nuclei and electrons $\hat{V}_{coulomb}$, and the electronic Hamiltonian $\hat{H}_{elec}(r, R)$:

$$\hat{H} = \hat{T}_{nuc}(R) + \hat{V}_{coulomb} + \hat{H}_{elec}(r, R) \tag{10}$$

where r and R refer to the position of the electrons and nucleus respectively. It has been demonstrated in extensive research that a direct solution of the Schrödinger equation is computationally impossible for large scale systems. Thus, DFT instead looks at the electron density of the many-body system, where a set of single-electron Schrödinger-like equations are solved, known as the Kohn-Sham equations (Gilbert, 1975; Jensen, 2017). Because the ground state energy of the system can be identified as a unique functional of electron density, by assuming a valid energy functional and minimizing the overall system energy, the true electron density can be obtained. However, this energy functional contains the exchanges and correlations of the system, which is not known except for the free electron gas (Gritsenko et al., 1997). Thus, a highly-accurate energy functional would be important to the accuracy of the DFT calculation.

In this work, a clean hydroxylated HfO_2 (111) surface is chosen as the starting point of the deposition. In order to compute the correct kinetic parameters of the reactions happening on this surface, an accurate HfO_2 (111) surface slab must be constructed. First, the monoclinic HfO_2 (space group $P2_1/c$) lattice structure is built based on the repetition of its unit cell. Then, Gaussian09 is used to optimize this unit cell with the built-in periodic boundary condition (PBC) calculation (Frisch et al., 2000). The resulting structure, which is shown in Fig. 2(a), is then imported into the Python CatKit Package, and the (111) surface slab, which is shown in Fig. 2(b) is calculated and generated according to the optimized lattice parameter using the CatGen module. The resulting slab is then terminated with hydroxyl groups, and the surface layer is relaxed while the atom positions in the bulk are fixed. Additionally, a single-$Hf(NCH_3)_2$-terminated HfO_2 slab is generated, as shown in Fig. 5(a). A molecular oxygen gas, O_2, is then carefully placed onto the terminating -NCH_3 group, as shown in Fig. 5(b), to simulate the O_2 desorption process and study the associated kinetics. The O_2-physisorbed slab is optimized to transitional state using the Hartree-Fock (HF) method with the Los Alamos National Laboratory 2-double-z (LanL2DZ) basis set to obtain an initial guess of the structure and the kinetics parameter for desorption (Hay and Wadt, 1985a,b; Wadt and Hay, 1985). Although it is ideal to compute the exact profile and the transitional state energy using the B3LYP functional, the large amount of atoms causes the computation to be very slow. Therefore, based on the initial guess, we perform trial-and-error to obtain the kinetics parameter for the O_2 desorption reaction that leads to reasonable simulation results as shown in the following section.

The important DFT-computed activation energies for Hf-cycle, which are derived from Shirazi and Elliott (2014), are listed as follows: $r_{Hf,des} = 23.29$ kcal/mole, $r_{Hf_a,che1} = 20.52$ kcal/mole, and $r_{Hf_b,che2} = 24.67$ kcal/mole. The sticking coefficient of TDMAHf, which is critical to the physisorption rate, is discussed in Dorsmann and Kleijn (2007). In addition, the kinetics of the barrierless plasma species and the oxida-

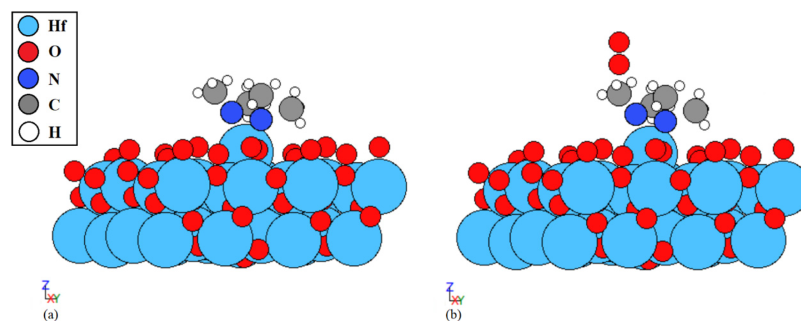


Fig. 5 – (a) Single-$Hf(NCH_3)_2$-terminated HfO_2 2 × 2 ×1 slab, and (b) with O_2-physisorbed.

tion energy of neutral oxygen, $r_{3_{O_2,oxidation}} = 58.48$ kcal/mole, are described in detail by Jeon and Won (2008). The desorption energy, $r_{3_{O_2,des}} = 56.01$ kcal/mole and the structural hindrance information are calculated and estimated in this paper.

2.6. Simulation model result and validation

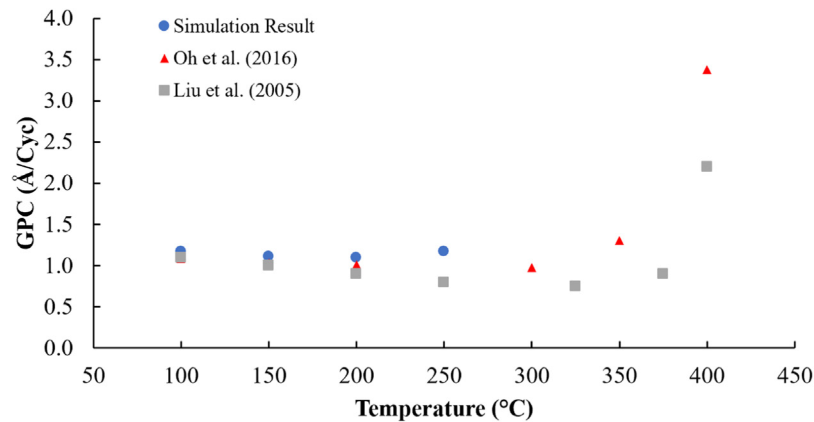
In order to validate the proposed kMC model, we compare the model output with experimental and simulation results reported in the literature (Choi et al., 2004; Liu et al., 2005; Lao et al., 2005; Park et al., 2006; Heil et al., 2007; Joo and Rossnagel, 2009; Provine et al., 2016; Oh et al., 2016). Specifically, we will compare the exact growth per cycle (GPC) values under several operating conditions from the simulation model with the individual GPC reported from past works. In addition, the effect of individual operating parameter on the GPC is studied, where the major factors we consider are the substrate temperature and the precursor partial pressure of each half-cycle. It is noteworthy that the partial pressure development depends on the gas-phase momentum and species transports in the main reactor chamber, which are heavily influenced by the reactor's physical length scale and detailed design. In particular, there will transiently be a radial distribution of neutral and plasma species on the substrate surface. Nevertheless, a well-developed flow will lead to a rather uniform profile based on the input operating condition. As a result, we will postulate a reasonable surface condition and investigate the exact influence of the gas-phase profile in future work. Moreover, although the proposed kMC model does not directly report the GPC as the experimental probing methods like ellipsometry and X-ray photoelectron spectroscopy (XPS), the lattice constants and the coverage information can be used to estimate the simulation growth rate. Base on the layer thickness of 2.7 Å, investigated from the slab optimization, the GPC is inferred with the accumulated coverage from the kMC simulation. Similar to the majority of experiments, the simulation model reports an overall linear growth with respect to increasing cycle numbers (Liu et al., 2005; Provine et al., 2016).

The GPC from the model is first compared with a variety of operating conditions for PEALD and ALD tested in literature, as shown in Table 1. The experimental condition is characterized by the substrate temperature and the precursor dosage, which is a function of precursor partial pressure and the half-cycle or pulse time. Choi et al. (2004), as shown in Table 1[1], uses a substrate temperature of 250 °C and a long precursor pulse time of 7 s with a relatively high reactor pressure of 1 torr, and Park et al. (2006), as shown in Table 1[2], reports an even higher reactor pressure of 3 torr but with a slightly lower half-cycle time of 2 s. Under these conditions, preliminary calculations show that the supply is sufficient for the PEALD system and nearly sufficient for the corresponding ALD system. The above literature report PEALD GPC from 1.06 to 1.3 Å and ALD GPC of 1.0 Å, which are similar to the model output, where a steady-steady GPC of around 1.17 Å is reached for PEALD and a slower growth can be reached for ALD, depending on the exact boundary condition. In addition, Provine et al. (2016), as shown in Table 1[3], investigates a relatively lower substrate temperature of 200 °C, 0.25 s of precursor pulse, 30 sccm flow of oxygen gas and a low flow rate of carrier gas, which can be characterized as a low overall dosage. It reports a slightly lower GPC for PEALD and a substantially low GPC for ALD. This operating condition is tested with our simulation model, which shows that a limited dosage and reduced temperature have a huge influence on the deposition contributed by neutral O_2 species. Moreover, Lao et al. (2005), as shown in Table 1[4], focuses on the room temperature PEALD on a small scale reactor with sufficient precursor dosage and reports a GPC of 0.8–1.1 Å under different probing methods. The room temperature PEALD is also tested in our simulation model, where we demonstrate that a similar GPC for PEALD can be reproduced and there is nearly no deposition from the neutral O_2 gas.

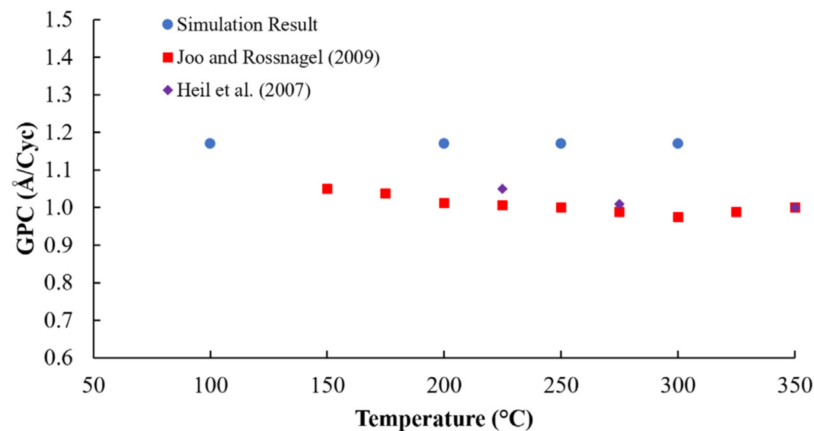
In addition to the GPC under literature-reported experimental conditions, we also validate our model with the effects of changing operating parameters on the deposition rate. As discussed in the previous paragraph, the substrate temperature is one of the most critical parameters in determining the GPC. However, the substrate temperature can be tricky to manipulate because it may have different effects on the two half-cycles due to different kinetic parameters. First, the influence of substrate temperature on the Hf-Cycle is shown in Fig. 6(a), and is compared with Oh et al. (2016) and Liu et al. (2005), labeled as red upper triangles and grey squares, respectively. Both works focus on the Hf-Cycle from the thermal ALD with TDMAHf as the precursor and demonstrate that the temperature has a significant influence on Hf-Cycle. Similar to our model, labeled by the blue dot, the film deposition rate is slightly decreasing in the lower temperature range and has a rebounding trend in the higher temperature range, because the adsorption complex is stabilized under low temperature. The decreasing of GPC in the lower temperature range is more significant according to the data reported by Liu et al. (2005) because the experimental film growth includes the impurity from carbon and hydrogen. Those impurities may have come from side reactions that are not considered in the kMC model reaction set. As the temperature increases, the impurity content gradually decreases and causes the apparent growth rate to decrease. Moreover, it can be seen that there is a dramatic increase in the film thickness, because of an undesirable

Table 1 – Comparison of GPC between literature values and simulation results under various operating conditions.

Condition	250 °C High dosage		200 °C Low dosage		Room temperature	
Method	[1] & [2]	Simulation	[3]	Simulation	[4]	Simulation
PEALD	1.06–1.3	~1.17	1.0–1.1	<1.17	0.8–1.1	1.1
ALD	1.0	<1.17	0.12–0.17	<0.12	N/A	Low



(a)



(b)

Fig. 6 – Substrate temperature influence on GPC: (a) influence on the Hf-cycle, and (b) influence on the O-Cycle.

rapid decomposition in the temperature range of 275–300 °C as demonstrated in the FT-IR analysis. Our simulation does not identify this decomposition mechanism because it is a significant variation from the standard deposition reaction set and is known to yield undesirable products. On the other hand, the effect on the plasma half-cycle (O-Cycle) is shown in Fig. 6(b). Heil et al. (2007) shows that the growth rate is fairly constant in the region of 230–350 °C, as shown by the purple diamond. Similarly, Joo and Rossnagel (2009) has developed a simulation model on the O_2 plasma and validate the plasma's influence on a small scale process, which also shows a steady GPC, labeled by the red squares. A comparable trend is reflected in our simulation. We consider a scenario where a sufficiently supplied Hf-Cycle is followed by a limited plasma dosage (time × pressure), and the result shows that the temperature does not have a drastic influence on the GPC, as shown by the blue dot. Additionally, the effect of excessive plasma pulse is investigated by Heil et al. (2007) and Provine et al. (2016), where no increasing GPC is observed from the extended pulsing time. Our simulation has demonstrated a similar result, in which the surface oscillates around steady states after enough dosage is given.

3. Machine learning

The half-cycle time, defined to be the time need to reach the self-limiting steady-state under each precursor pulse, is one of the most important considerations for the industrial thin-film production. According to the developed kMC model, the two half-cycles have a complicated interaction with each other on the structural detail like the existence of steric hindrance and the competition between available reaction cites. Therefore, the input–output relationship between half-cycle time and experimental conditions cannot be easily characterized by a small set of experiments. However, the cost of the research and development limits the exploration of a wide range of operating conditions for optimal throughput. Let alone the cost for energy and equipment maintenance, the precursor, TDMAHf, for a 100-cycle experiment of the industrial size wafer could

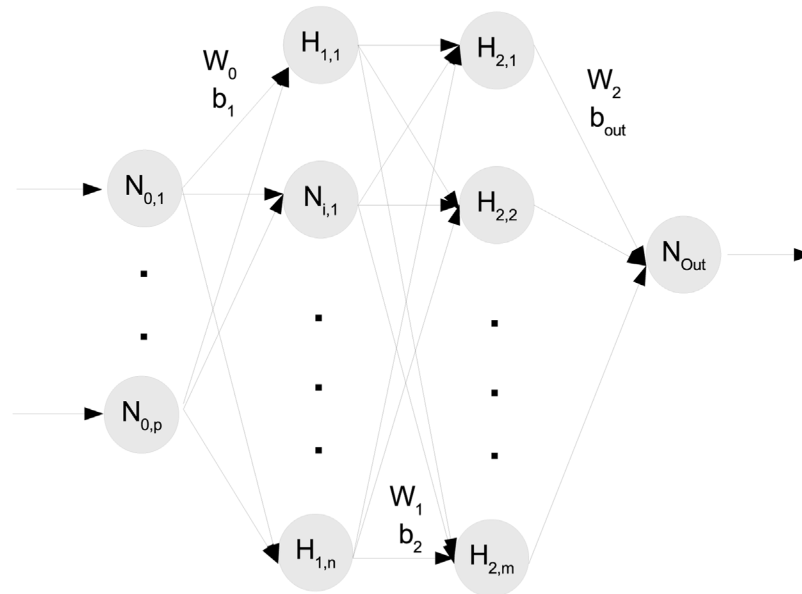


Fig. 7 – Illustration of the densely connected NN.

cost more than 1000 USD. On the other hand, the developed microscopic kMC model can provide cost-effective information but is still computationally expensive to be implemented in real-time and does not have a closed-form solution, as introduced in Section 1. As a result, it is crucial to collect a database and develop an efficient neural network (NN) model that estimates the half-cycle time needed, given a variety of input conditions. The objective of our model is to both investigate a wide range of operating conditions, which is required to predict the boundary of the feasible region and capture the precise input–output relationship, which is needed to determine the optimal operating condition. Thus, the desired model needs to pursue both good predicting precision and generality. Nevertheless, the traditional NNs with backpropagation often struggle with model accuracy and over-fitting. Thus, to successfully achieve both two aims, the Bayesian regularized artificial neural network (BRANN) is chosen as the solution. With the addition of the Bayesian inference to the standard ANN, BRANN becomes a more robust network for precise regression because the Bayesian regularization algorithm reformulates the complex non-linear regression to a statistically well-posed problem. Also, by incorporating Occam's razor principle to turn off the non-relevant weight in the model training efficiently, the BRANN network penalizes the excessive complexity and resolves the over-training situation (Burden and Dave, 2008).

3.1. *Model construction*

The construction of the BRANN network is implemented in the MATLAB machine learning toolbox. A dense connection is chosen, where all neurons in the neighbor layers are connected. A two-layer structure example is shown in Fig. 7, and the mathematical formulation of the neural network is shown as follows:

$$N_{1,j} = f_0\left(\sum_{i=1}^{p} W_{0,ij}N_{0,i} + B_{1,j}\right) \tag{11}$$

$$N_{2,j} = f_2\left(\sum_{i=1}^{n} W_{1,ij}N_{1,i} + B_{2,j}\right) \tag{12}$$

$$N_{out,j} = f_3\left(\sum_{i=1}^{m} W_{2,ij}N_{2,i} + B_{out,j}\right) \tag{13}$$

where each equation describes the mathematical relationship between the connected layers. *N* denotes the neuron, *W* denotes the weight matrix, and *B* denotes the bias matrix. The subscript notation 0 refers to the input layer, 1 and 2 refer to the first and second hidden layer, respectively, and *out* refers to the output layer. In each equation, *i* and *j* refer to the index of neurons in the previous and the subsequent layer, and *ij* refers to the matrix index in the weight matrix relating to *i* and *j*. *p*, *n*, and *m* describe the total number of neurons in the input, first and the second layer, respectively. In addition to the linear combination, activation function *f* is added to allow for the NN to describe the non-linearity. In the BRANN model, instead of the usual Rectified Linear Unit (ReLU) function, the hyperbolic tangent sigmoid function (tansig) is used as the transfer function:

$$tansig(x) = \frac{2}{1 + \exp(-2x)} - 1 \tag{14}$$

where *x* is the value passed in the network. This function has a range of [0, 1], which makes it the right candidate to fit functions that map into probability spaces, allowing the network to interpret the probability of an event in a more intuitive manner.

The goal of the training is to minimize the cost function, S(w), with respect to the parameters in the aforementioned structure. Unlike the traditional mean square error (MSE) function, which is typically used for regression application, the

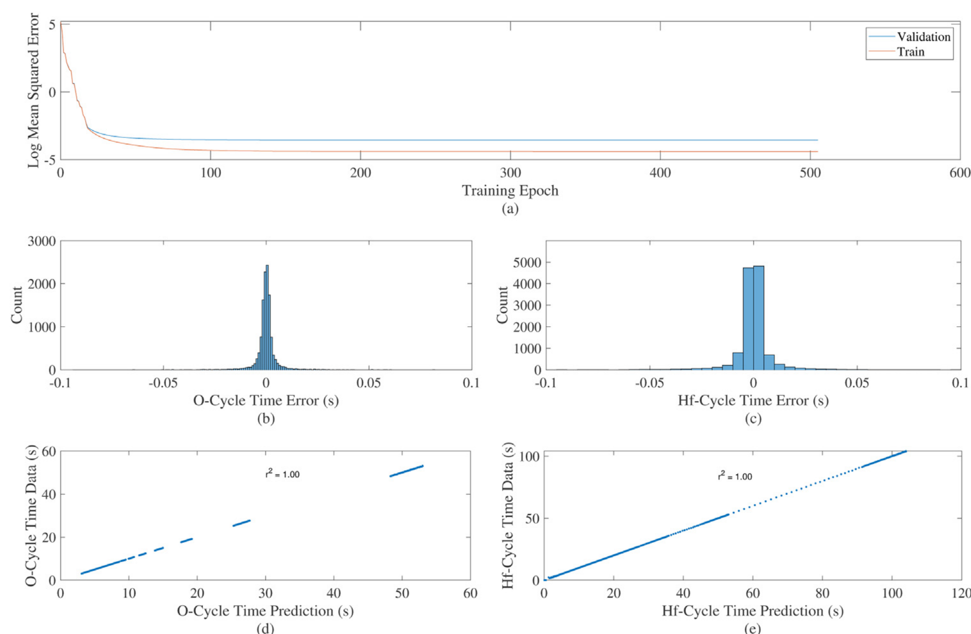


Fig. 8 – Performance of the NN: (a) Training and validation error over the training process. (b) Prediction error distribution histogram for O-Cycle. (c) Prediction error distribution histogram for Hf-Cycle. (d), (e) Correlation accuracy of the predicted time and the actual time for O-Cycle and Hf-Cycle, respectively, where the x-axis is the predicted time to reach steady-state from the neural network and the y-axis is the actual time to reach steady-state from the kMC model.

BRANN network adopts additional hyperparameters to regularize the cost function as follows:

$$S(w) = \beta \sum_{i=1}^{N_D} (y_i - f(x_i, w))^2 + \alpha \sum_{j=1}^{N_w} w_j^2 \tag{15}$$

where i is the index for the data sample, N_D is the total number of training data, y_i is the desired output value of ith entry, and $f(x_i, w)$ is the predicted value dependent on w and the input x_i. w is the weight vector, which is initialized through Gaussian distribution, j is the index for the weight, and N_w is the total number of weight parameters. α and β are the hyperparameters for the weight regularization, initialized from a uniform random distribution. In order to arrive at the optimal weight vector, the maximum a posteriori (MAP) method is solved with the Levenberg–Marquardt algorithm (LMA) (Moré, 1978). Similar to other numerical optimization algorithms, the LMA is an iterative procedure. It behaves like a gradient-descent method when the parameters are far from the optimal value, while more similar to the Gauss–Newton method when the parameters are close to optimal. Compared to the RMSProps and gradient descent method, LMA is demonstrated to have a more stable convergence progressing toward the global minimum (Bazzi et al., 2018). The original backpropagation is modified with the addition of Bayes' theorem to approximate the exact Bayesian inference of the posterior probability, which leads the network to faster regularization and eliminates the need for dropout:

$$P(w|D) = \frac{P(w)P(D|w)}{P(D)} \tag{16}$$

where D refers to the dataset. In the meantime, the hyperparameters also need to be iteratively updated during each epoch according to Bayesian inference. This dual step optimization is repeated until the maximum epoch self-consistency is achieved. The detailed network construction and theory explanation are discussed in Burden and Dave (2008) and MacKay (1992).

Table 2 – Temperature, pressure, predicted total time and conventional total time for process optimization. Predicted total time is also corrected with the one-second gas-phase transport time.

#	Temperature (K)	Pressure (Pa)	Predicted total time (s)	Original total time (s)
1	300	20	11.20	10.0
2	330	25	10.05	10.0
3	360	30	9.28	10.0
4	390	35	8.71	10.0
5	420	40	8.37	10.0

3.2. *Dataset generation*

In advance of the model training, a database should be generated, which contains the respective half-cycle times corresponding to a reasonable range of input operating conditions. Therefore, we have developed an automated workflow to run the kMC model for the two half-cycles and to test a variety of operating condition combinations. As shown in Fig. 1, the top box specifies the kMC model workflow. The simulation starts at $t = 0$ where the operating temperature and partial pressure for the Hf-Cycle are given to the kMC model. The event selection scheme is executed and the time progression is calculated until the model reaches the termination criterion, where the surface is saturated under the restriction from steric hindrance. Then, the structural detail from the Hf-Cycle is passed to the O-Cycle, and the partial pressures for the O-Cycle precursors are given to the model as inputs. This workflow could be repeated until the desired cycle number is achieved. The overall time needed is stored as the final result. The automated workflow is utilized to explore a broad range of operating condition combinations. To boost the computational speed and efficiently utilize all available

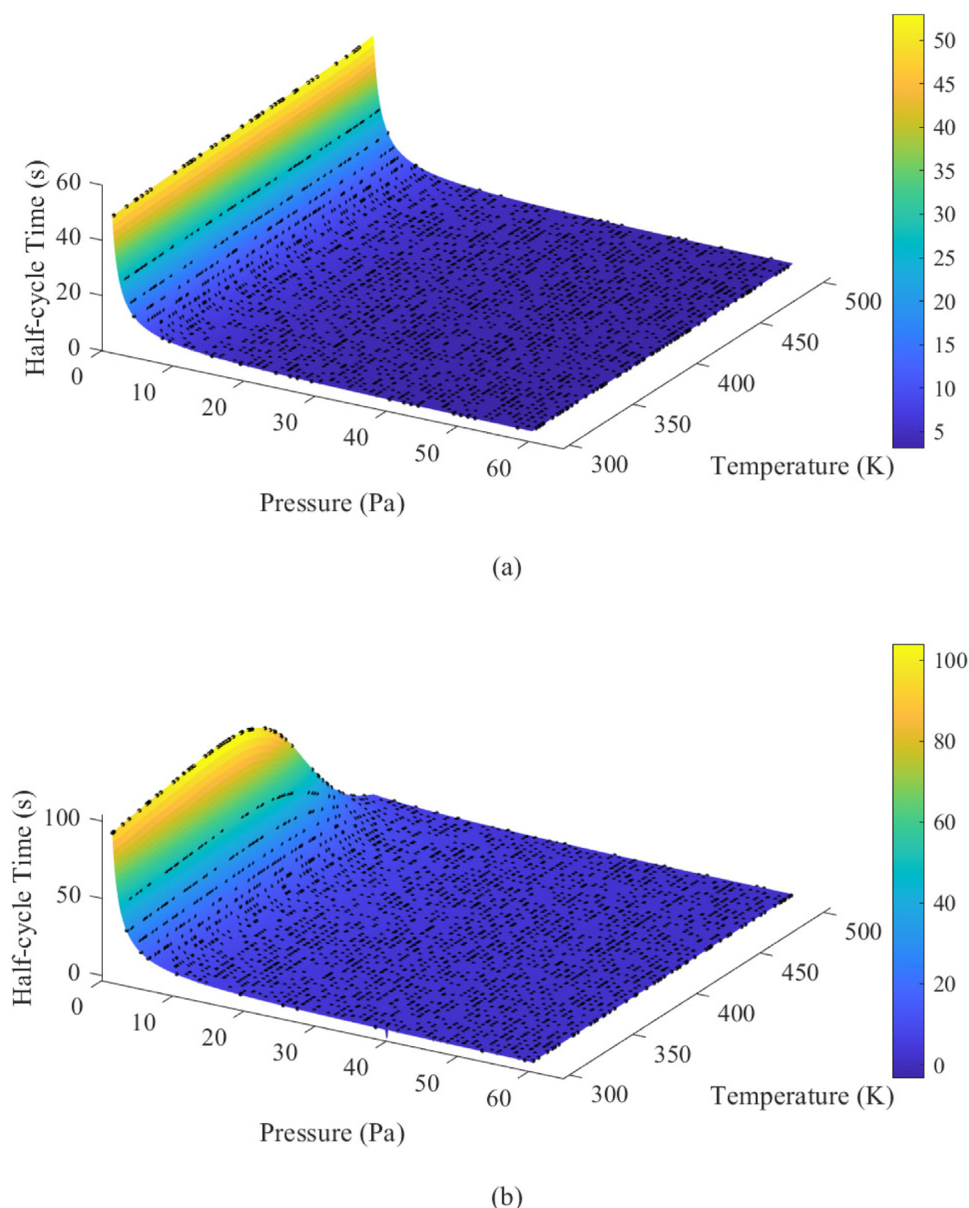


Fig. 9 – Neural network fitting results and original training data. *x*-axis is pressure (Pa). *y*-axis is temperature (K). *z*-axis is the half-cycle time from prediction and training data. Black dots denote the training data, and the colormap denotes the magnitude of predicted time from the neural network. (a) NN fitting result for O-Cycle. (b) NN fitting result for Hf-Cycle.

computational resources, we adopt a distributed computation using the UCLA Hoffman2 cluster. With the developed workflow, the kMC models are submitted as computation jobs to the Hoffman2 job queue, and the job scheduler will distribute the jobs onto different cores to allow parallel runs of the kMC models. Because of the competition of the reverse and forward reaction and the stochastic nature of kMC, the final lattice condition may vary under different input combinations and keep fluctuating. Nevertheless, the overall coverage will converge to steady-state, where the film structure does not differ dramatically. Thus, for the NN characterization, the half-cycle time needed for the kMC model to reach 99.7% of the final coverage is taken to minimize the noise associated with the fluctuation. The results are collected and compiled automatically. In this work, a single cycle simulation is investigated under a set of operating conditions where the substrate temperature ranges from 298 to 503 K, and the overall pressure ranges from 1 to 60 Pa, with an interval of 1 K and 1 Pa, respectively. The available computation resources allow us to run 500 jobs actively in parallel and queue over 4000 jobs automatically for execution. The full database can be generated under two weeks. In addition, the database is further randomly divided into training, validation, and testing dataset with a ratio of 7:1.5:1.5. The training and validation datasets are applied in the model parameter updating and evaluation, while the testing dataset is used to assess the final model performance.

3.3. *NN result and cycle time optimization*

The numbers of neurons and layers are optimized via a grid search (Svozil et al., 1997). In particular, the input layer consists of two neurons, which represent the surface temperature and the total precursor pressure, respectively. The output layer consists of two neurons, which represent each of the half-cycle completion times to reach full coverage. Two hidden layers are included, which consist of 20 and 15 neurons, respectively. Good training and validation accuracy have been achieved as shown in Fig. 8 (a), in which we can see that the training error

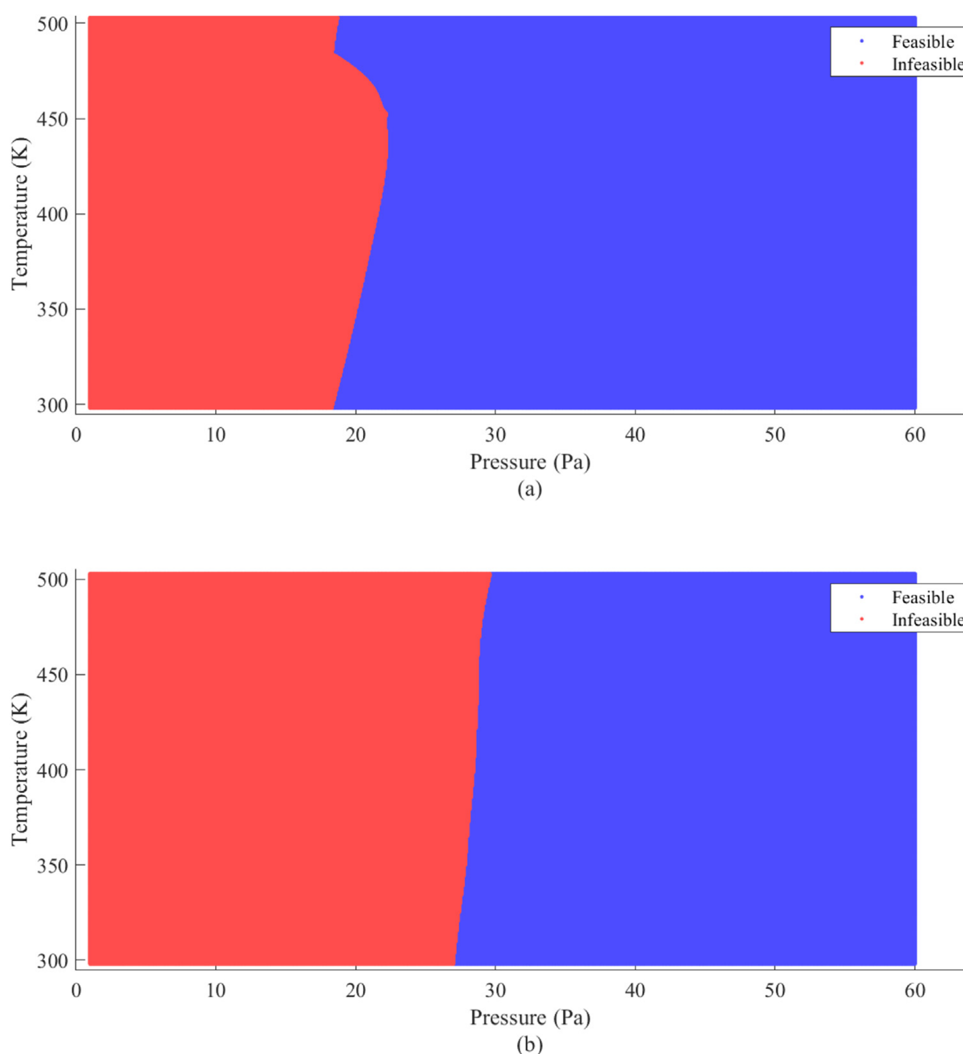


Fig. 10 – Operational feasibility of HfO_2 under fixed operating conditions. The blue region denotes the operating conditions where both half-cycles can be completed under desired half-cycle time, whereas the red region denotes otherwise. (a) Effective half-cycle time = 5 s, without considering gas-phase transport. (b) Effective half-cycle time = 4 s, accounting for 1 s of gas-phase transport.

quickly drops and the validation mean square converges to $2.84 \times 10^{-4}\, s^2$. The fluctuation at the beginning of the training could be attributed to the stochastic nature of Bayesian inference used by BRANN. As the training progress moves towards convergence, the curve is smoother because the parameters are close to optimal values. The testing dataset is compared with the model prediction, where the NN achieves a mean absolute error of 0.017 s. Such a low testing error demonstrates a good model accuracy because it is already in the range of steady-state variation of the kMC runs. In Fig. 8(b) and (c), the error distribution histograms demonstrate that, for both half-cycles, the deviation of the prediction from the actual half-cycle completion highly resemble the normal distribution, which demonstrates that the NN fitting is unbiased. Additionally, in Fig. 8(d) and (e), the R-squared correlations between the kMC simulation results and the predicted time are 1.00 for both half-cycles, which demonstrates that the fitting result is in good agreement with the original kMC output. This fitting result can be further visualized in Fig. 9. The grid search is completed on two computers with 64G RAM and 8-cores in one day. After the NN is trained, it takes less than one second to perform the prediction using the same computer.

With the trained BRANN model, we can explore the operating conditions to optimize the half-cycle time. An extended half-cycle time is not ideal for the PEALD process as it significantly reduces the industrial throughput and incurs unnecessary operating costs. Currently, appropriate operating conditions are investigated through trial and error for each reactor design. For example, various works have investigated the O-Cycle pulse time ranging from 1 to 60 s and the substrate temperature ranging from room temperature to 375 °C. Thus, a feasible operating range is valuable to be determined so that lab works can be completed with a reasonable experimental cost and effort. A five-second time frame for each half cycle is chosen as the reference half-cycle time, which is the typical production time of a thin-film of the modern gate oxide. By using the developed BRANN model, as shown in Fig. 10(a), we can identify the feasible range of operating conditions where both cycles can finish under 5 s. However, as discussed in Section 2.6, the time for gas-phase transport should also be accounted for. Our preliminary study has shown that a radial distributed gas-phase transport will lead to a typical development time ranges from 1 to 1.5 s for a standard remote PEALD reactor design. This rather long time range makes sure the prevention of potential damage onto the substrate surface through the high-energy plasma sheathing and high-speed perpendicular flow. The BRANN model is also utilized to estimate the feasible operating region by subtract-

ing the gas-phase transport time from the effective half-cycle time. As shown in Fig. 10(b), when 1 s of gas-phase transport time is accounted for, there is a significant change and reduction of the feasible operating range. The curvature of boundary shown in Figure 10 is largely due to the nature of the kinetics. Since our kMC model modifies the standard kMC algorithm to include transient concentrations as weights, the direct relationship between the boundary and the rate of each reaction is not very straightforward. However, the nonlinearity in the boundary could provide us some insight into the overall trend of reaction rates at certain operating conditions. For example, at $P = 18$ Pa and $T = 480$ K, due to the increase in temperature, the desorption rate would increase to a point where it surpasses the increase of the forward deposition reaction rates. This overall trend shift would cause the feasibility region to shrink and lead to the kink at the boundary. Additionally, the narrow P-zone is due to the consideration of the operation of PEALD under fixed temperature and pressure in this work. In realistic operations, low P-zone would typically only exist a short amount of time, therefore it would not lead to the final incompletion.

Finally, we utilize the developed BRANN model to perform process optimization. As demonstrated in Table 2, a series of operations under fixed operating conditions are examined, and the predicted total time for both cycles are compared with the conventional fixed total time of 10 s. The precursor partial pressure is limited under 40 Pa to reduce the precursor cost and the substrate temperature is limited under 150 °C (423 K) to simulate a low-temperature deposition scenario which is critical to the temperature-sensitive materials. For experimental conditions 1 and 2 in Table 2, if the conventional fixed time is used, the product will be uncompleted because insufficient time is given for the cycle. For experimental conditions 3, 4, and 5 in Table 2, by using the accurate prediction from the NN model, the operation time could be reduced by up to 16%, which corresponds to a significant enhancement in the industrial throughput and operating cost.

4. Conclusion

In this work, we developed a microscopic model of the PEALD of HfO_2 using TDMAHf and oxygen plasma as precursors. A specially tailored kMC algorithm is utilized, which is dependent on the DFT calculation and is validated by experimental results. The GPC reported from the simulation model deviates from the experimental results up to 6.3% under various operating conditions. Moreover, the model successfully characterized the influence of individual operating parameters on the GPC that are observed in the existing literature. After the simulation model has been validated, we constructed a BRANN model based on the database generated from the kMC model for a wide range of operating conditions. Using the BRANN model, we performed data analysis to determine the feasible operating region and to identify the optimal process condition, where a reduction of half-cycle time up to 16% can be achieved. Instead of using the trial and error approach, the kMC model provides an operating guideline for both industrial manufacturing and experimental researches. By using the optimal time predicted by the model, the production throughput can be greatly improved, along with a significant reduction for labor and operating costs. It is notable that the developed kMC scheme is not limited to the deposition of HfO_2. It could be extended to the deposition of any XY_2-type compound with similar reaction routes. Therefore, in order to implement a generic microscopic PEALD model for XY_2-type compound, we generalized our kMC scheme and implemented the model in Python to take advantage of its modularity and accessibility. Thus, the developed model formulated a general methodology to explore the precursor applicability, boost the PEALD throughput, and reduce the operating cost.

Conflict of interest

None declared.

Acknowledgments

Financial support from the National Science Foundation is gratefully acknowledged.

References

Armaou, A., Christofides, P.D., 1999. Plasma enhanced chemical vapor deposition: modeling and control. Chem. Eng. Sci. 54, 3305–3314.

Battaile, C.C., Srolovitz, D.J., 2002. Kinetic Monte Carlo simulation of chemical vapor deposition. Annu. Rev. Mater. Res. 32, 297–319.

Bazzi, T., Ismail, R., Zohdy, M., 2018. Comparative performance of several recent supervised learning algorithms. Int. J. Comput Inform. Technol. 7, 49–54.

Burden, F., Dave, W., 2008. Bayesian Regularization of Neural Networks, Artificial Neural Networks: Methods and Applications. Springer, pp. 23–42.

Choi, S., Koo, J., Jeon, H., Kim, Y., 2004. Plasma-enhanced atomic-layer deposition of a HfO_2 gate dielectric. J. Kor. Phys. Soc. 44, 35–38.

Cortright, R.D., Dumesic, J.A., 2001. Kinetics of heterogeneous catalytic reactions: analysis of reaction schemes. Adv. Catal. 46, 161–264.

Crose, M., Zhang, W., Tran, A., Christofides, P.D., 2018. Multiscale three-dimensional CFD modeling for PECVD of amorphous silicon thin films. Comput. Chem. Eng. 113, 184–195.

Ding, Y., Zhang, Y., Kim, K., Tran, A., Wu, Z., Christofides, P.D., 2019a. Microscopic modeling and optimal operation of thermal atomic layer deposition. Chem. Eng. Res. Des. 145, 159–172.

Ding, Y., Zhang, Y., Ren, Y.M., Orkoulas, G., Christofides, P.D., 2019b. Machine learning-based modeling and operation for ALD of SiO_2 thin-films using data from a multiscale CFD simulation. Chem. Eng. Res. Des. 151, 131–145.

Dkhissi, A., Estève, A., Mastail, C., Olivier, S., Mazaleyrat, G., Jeloaica, L., Djafari-Rouhani, M., 2008. Multiscale modeling of the atomic layer deposition of HfO_2 thin film grown on silicon: How to deal with a kinetic Monte Carlo procedure. J. Chem. Theory Comput. 4, 1915–1927.

Dorsmann, R., Kleijn, C.R., 2007. A general correction to surface reaction models based on reactive sticking coefficients. Chem. Vapor Depos. 13, 91–97.

Elliott, S.D., Greer, J.C., 2004. Simulating the atomic layer deposition of alumina from first principles. J. Mater. Chem. 14, 3246–3250.

Epling, W.S., Peden, C.H., Henderson, M.A., Diebold, U., 1998. Evidence for oxygen adatoms on TiO_2 (110) resulting from O_2 dissociation at vacancy sites. Surf. Sci. 412, 333–343.

Frisch, A., Nielsen, A.B., Holder, A.J., 2000. Gaussview User Manual. Gaussian Inc., Pittsburgh, PA, pp. 556.

George, S.M., 2009. Atomic layer deposition: an overview. Chem. Rev. 110, 111–131.

Gilbert, T., 1975. Hohenberg–Kohn theorem for nonlocal external potentials. Phys. Rev. B 12, 2111.

Gritsenko, O.V., Schipper, P.R.T., Baerends, E.J., 1997. Exchange and correlation energy in density functional theory:

comparison of accurate density functional theory quantities with traditional Hartree–Fock based ones and generalized gradient approximations for the molecules Li_2, N_2, F_2. J. Chem. Phys. 107, 5007–5015.
Han, B., Zhang, Q., Wu, J., Han, B., Karwacki, E.J., Derecskei, A., Xiao, M., Lei, X., O'Neill, M.L., Cheng, H., 2011. On the mechanisms of SiO_2 thin-film growth by the full atomic layer deposition process using bis(t-butylamino)silane on the hydroxylated SiO_2(001) surface. J. Phys. Chem. C 116, 947–952.
Hay, P.J., Wadt, W.R., 1985a. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. J. Chem. Phys. 82, 299–310.
Hay, P.J., Wadt, W.R., 1985b. Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. J. Chem. Phys. 82, 270–283.
Heil, S., Van Hemmen, J., Hodson, C., Singh, N., Klootwijk, J., Roozeboom, F., Van de Sanden, M., Kessels, W., 2007. Deposition of TiN and HfO_2 in a commercial 200 mm remote plasma atomic layer deposition reactor. J. Vacuum Sci. Technol. A Vacuum Surf. Films 25, 1357–1366.
Huang, J., Hu, G., Orkoulas, G., Christofides, P.D., 2010. Dependence of film surface roughness and slope on surface migration and lattice size in thin film deposition processes. Chem. Eng. Sci. 65, 6101–6111.
Ishikawa, K., Karahashi, K., Ichiki, T., Chang, J.P., George, S.M., Kessels, W., Lee, H.J., Tinck, S., Um, J.H., Kinoshita, K., 2017. Progress and prospects in nanoscale dry processes: how can we control atomic layer reactions? Jap. J. Appl. Phys. 56 (06), HA02.
Jensen, F., 2017. Introduction to Computational Chemistry. John Wiley & Sons.
Jeon, H., Won, Y., 2008. The reaction pathways of the oxygen plasma pulse in the hafnium oxide atomic layer deposition process. Appl. Phys. Lett. 93, 124104.
Jeong, S.J., Gu, Y., Heo, J., Yang, J., Lee, C.S., Lee, M.H., Lee, Y., Kim, H., Park, S., Hwang, S., 2016. Thickness scaling of atomic-layer-deposited HfO_2 films and their application to wafer-scale graphene tunnelling transistors. Sci. Rep. 6, 20907.
Joo, J., Rossnagel, S.M., 2009. Plasma modeling of a PEALD system for the deposition of TiO_2 and HfO_2. J. Kor. Phys. Soc. 54, 1048.
Kanomata, K., Pansila, P., Ahmmad, B., Kubota, S., Hirahara, K., Hirose, F., 2014. Infrared study on room-temperature atomic layer deposition of TiO_2 using tetrakis (dimethylamino) titanium and remote-plasma-excited water vapor. Appl. Surf. Sci. 308, 328–332.
Kimaev, G., Ricardez-Sandoval, L.A., 2019. Nonlinear model predictive control of a multiscale thin film deposition process using Artificial Neural Networks. Chem. Eng. Sci. 207, 1230–1245.
Knoops, H.C.M., Langereis, E., Van De Sanden, M.C.M., Kessels, W.M.M., 2010. Conformality of plasma-assisted ALD: physical processes and modeling. J. Electrochem. Soc. 157, G241–G249.
Kukli, K., Ritala, M., Schuisky, M., Leskelä, M., Sajavaara, T., Keinonen, J., Uustare, T., Hårsta, A., 2000. Atomic layer deposition of titanium oxide from TiI_4 and H_2O_2. Chem. Vapor Depos. 6, 303–310.
Kumar, N., George, B.P.A., Abrahamse, H., Parashar, V., Ray, S.S., Ngila, J.C., 2017. A novel approach to low-temperature synthesis of cubic HfO_2 nanostructures and their cytotoxicity. Sci. Rep. 7, 1–14.
Lao, S.X., Martin, R.M., Chang, J.P., 2005. Plasma enhanced atomic layer deposition of HfO_2 and ZrO_2 high-k thin films. J. Vacuum Sci. Technol. A: Vacuum Surf. Films 23, 488–496.
Lee, C.S., Oh, M.S., Park, H.S., 2006. Plasma enhanced atomic layer deposition (PEALD) equipment and method of forming a conducting thin film using the same thereof. US Patent 7,138,336.
Liu, X., Ramanathan, S., Longdergan, A., Srivastava, A., Lee, E., Seidel, T.E., Barton, J.T., Pang, D., Gordon, R.G., 2005. ALD of hafnium oxide thin films from tetrakis (ethylmethylamino) hafnium and ozone. J. Electrochem. Soc. 152, G213–G219.
Liu, X., Ramanathan, S., Seidel, T.E., 2003. Atomic layer deposition of hafnium oxide thin films from tetrakis (dimethylamino) hafnium (TDMAH) and ozone. MRS Online Proc. Library Arch., 765.
Liu, Y., Liu, H., 2018. Development of 3D polymer DFT and its application to molecular transport through a surfactant-covered interface. AIChE J. 64, 238–249.
MacKay, D.J.C., 1992. Bayesian interpolation. Neural Comput. 4, 415–447.
Melo, L., Burton, G., Kubik, P., Wild, P., 2016. Long period gratings coated with hafnium oxide by plasma-enhanced atomic layer deposition for refractive index measurements. Opt. Express 24, 7654–7669.
Moré, J.J., 1978. The Levenberg–Marquardt algorithm: implementation and theory. In: Numerical Analysis. Springer, pp. 105–116.
Nalwa, H. (Ed.), 2002. Handbook of Thin Films, Vol. 1. Academic Press, Burlington.
Oh, N.K., Kim, J.T., Ahn, J.K., Kang, G., Kim, S.Y., Yun, J.Y., 2016. The effects of thermal decomposition of tetrakis-ethylmethylaminohafnium (TEMAHf) precursors on HfO_2 film growth using atomic layer deposition. Appl. Sci. Converg. Technol. 25, 56–60.
Park, P.K., Roh, J.S., Choi, B.H., Kang, S.W., 2006. Interfacial layer properties of HfO_2 films formed by plasma-enhanced atomic layer deposition on silicon. Electrochem. Solid-State Lett. 9, F34–F37.
Pittal, S., Snyder, P.G., Ianno, N.J., 1993. Ellipsometry study of non-uniform lateral growth of ZnO thin films. Thin Solid Films 233, 286–288.
Provine, J., Schindler, P., Torgersen, J., Kim, H.J., Karnthaler, H.P., Prinz, F.B., 2016. Atomic layer deposition by reaction of molecular oxygen with tetrakisdimethylamido-metal precursors. J. Vacuum Sci. Technol. A: Vacuum Surf. Films 34 (01), A138.
Rai, V.R., Vandalon, V., Agarwal, S., 2010. Surface reaction mechanisms during ozone and oxygen plasma assisted atomic layer deposition of aluminum oxide. Langmuir 26, 13732–13735.
Rasoulian, S., Ricardez-Sandoval, L.A., 2015. Robust multivariable estimation and control in an epitaxial thin film growth process under uncertainty. J. Process Control 34, 70–81.
Rey, J.C., Cheng, L., McVittie, J.P., Saraswat, K.C., 1991. Monte Carlo low pressure deposition profile simulations. J. Vacuum Sci. Technol. A 9, 1083–1087.
Ribes, G., Mitard, J., Denais, M., Bruyere, S., Monsieur, F., Parthasarathy, C., Vincent, E., Ghibaudo, G., 2005. Review on high-k dielectrics reliability issues. IEEE Trans. Dev. Mater. Reliab. 5, 5–19.
Schuegraf, K., Abraham, M.C., Brand, A., Naik, M., Thakur, R., 2013. Semiconductor logic technology innovation to achieve sub-10 nm manufacturing. IEEE J. Electron Dev. Soc. 1, 66–75.
Schwille, M.C., Schössler, T., Barth, J., Knaut, M., Schön, F., Höchst, A., Oettel, M., Bartha, J., 2017. Experimental and simulation approach for process optimization of atomic layer deposited thin films in high aspect ratio 3D structures. J. Vacuum Sci. Technol. A: Vacuum Surf. Films 35 (01), B118.
Shi, X., Tielens, H., Takeoka, S., Nakabayashi, T., Nyns, L., Adelmann, C., Delabie, A., Schram, T., Ragnarsson, L., Schaekers, M., et al., 2011. Development of ALD $HfZrO_2$ with TDEAH/TDEAZ and H_2O. J. Electrochem. Soc. 158, H69–H74.
Shirazi, M., Elliott, S.D., 2014. Atomistic kinetic Monte Carlo study of atomic layer deposition derived from density functional theory. J. Comput. Chem. 35, 244–259.
Svozil, D., Kvasnicka, V., Pospichal, J., 1997. Introduction to multi-layer feed-forward Neural Networks. Chemometr. Intell. Lab. Syst. 39, 43–62.
Tinck, S., Bogaerts, A., 2011. Computer simulations of an oxygen inductively coupled plasma used for plasma-assisted atomic layer deposition. Plasma Sources Sci. Technol. 20, 015008.
Travis, C.D., Adomaitis, R.A., 2013. Modeling ALD surface reaction and process dynamics using absolute reaction rate theory. Chem. Vapor Depos. 19, 4–14.

Travis, C.D., Adomaitis, R.A., 2014. Modeling alumina atomic layer deposition reaction kinetics during the trimethylaluminum exposure. Theoret. Chem. Acc. 133, 1414.

Wadt, W.R., Hay, P.J., 1985. Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. J. Chem. Phys. 82, 284–298.

Won, Y., Park, S., Koo, J., Kim, S., Kim, J., Jeon, H., 2005. Initial reaction of hafnium oxide deposited by remote plasma atomic layer deposition method. Appl. Phys. Lett. 87, 262901.

Yun, S.J., Lim, J.W., Lee, J.H., 2004. PEALD of zirconium oxide using tetrakis (ethylmethylamino) zirconium and oxygen. Electrochem. Solid-State Lett. 7, F81–F84.

Zeng, Q., Oganov, A.R., Lyakhov, A.O., Xie, C., Zhang, X., Zhang, J., Zhu, Q., Wei, B., Grigorenko, I., Zhang, L., et al., 2014. Evolutionary search for new high-k dielectric materials: methodology and applications to hafnia-based oxides. Acta Crystallogr. Sect. C: Struct. Chem. 70, 76–84.

Zhang, Y., Ding, Y., Christofides, P.D., 2019. Multiscale computational fluid dynamics modeling of thermal atomic layer deposition with application to chamber design. Chem. Eng. Res. Des. 147, 529–544.

Zhang, Y., Ding, Y., Christofides, P.D., 2020. Integrating feedback control and run-to-run control in multi-wafer thermal atomic layer deposition of thin films. Processes 8, 18.