Digital Chemical Engineering 10 (2024) 100140

Contents lists available at ScienceDirect

Digital Chemical Engineering

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

Original article

Multiscale computational fluid dynamics modeling of an area-selective atomic layer deposition process using a discrete feed method

Henrik Wang [a], Matthew Tom [a], Feiyang Ou [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

Keywords:
Multiscale computational fluid dynamics modeling
Semiconductor manufacturing processes
Area-selective atomic layer deposition
Reactor optimal design
Reactor optimal operation

ABSTRACT

Area-selective atomic layer deposition (AS-ALD) is a beneficial procedure that facilitates self-alignment for transistor stacking by concentrating oxide growth on targeted areas of a substrate. However, AS-ALD is difficult to incorporate into semiconductor manufacturing industries due to difficulties such as minimal process data and a lack of insight into reactor design. To enable the industrial scale-up of AS-ALD, *in silico* modeling is necessary to characterize the process. Thus, this work proposes a multiscale computational fluid dynamics modeling framework that simultaneously describes the surface chemistry and ambient fluid behavior for an Al_2O_3/SiO_2 substrate. The multiscale model first involves *ab initio* molecular dynamics simulations to optimize molecular structures involved in the AS-ALD reactions. Next, a kinetic Monte Carlo simulation is performed to describe the stochastic surface chemistry behavior to determine the surface coverage, and deposition and byproduct rates. Lastly, computational fluid dynamics is performed to study the spatiotemporal behavior of the flow. The surface and flow field simulations are carried out in an integrated fashion. Various AS-ALD discrete feed reactor configurations with differing injection plate geometries were developed to investigate their impact on the processing time to achieve full surface coverage and film uniformity. Results indicate that the multi-inlet reactor model achieves minimal processing time while producing a high-quality film with the AS-ALD process.

1. Introduction

In a growing industrial world comprised of high-performance electronic devices, there exist many challenges in the production of semiconductor devices, which are vital to the innovation of modern-day electronics. The global dependency on semiconductors is appreciable in many industries, particularly so for smart technology, gaming and computing, biomedical technology (Chen and Chiu, 2022), and communication. However, this technological dependence has also spawned numerous challenges related to consistent manufacturing of high-performance semiconductor devices; in part, these challenges stem from the stringent design specifications that are common for modern semiconductor devices. Additionally, as the rudimentary projections of Moore's Law (Moore, 1998) for the densification of transistor materials predict further miniaturization of transistor length scales, chip production efficiency has reduced. Particularly, the continual reduction in transistor size has magnified short-channel effects, which degrade the computing power and power efficiency of transistors (Radamson, 2018). To resolve this issue, complex transistor designs, such as the gate-all-around (GAA), have been developed to address various design flaws including power and current losses, thereby optimizing the performance of the transistors (Zhang et al., 2012). However, transistor fabrication, especially at the nanoscale level, is difficult to implement in large-scale applications (Yuan and Sheng, 2010). Thus, there is a growing motivation to improve the accuracy and efficiency of the production method of these semiconducting materials. One method to achieve higher quality and productivity is to optimize the process by gathering a large amount of data by examining chemical processes within reactor models that enable process scale-up. This work will study the development of a three-dimensional reactor with a discrete feed mechanism for an area-selective atomic layer deposition process.

In atomic layer deposition (ALD) processes, precursor reagents are deposited onto the substrate surface to enable thin-layer growth in a bottom-up fabrication method (Zhang et al., 2022). Despite ALD being practical in most industrial applications, the process generally introduces alignment issues that are attributed to nonuniform surfaces and

* Corresponding author at: Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA, 90095-1592, USA.
E-mail address: pdc@seas.ucla.edu (P.D. Christofides).

https://doi.org/10.1016/j.dche.2024.100140
Received 21 December 2023; Received in revised form 16 January 2024; Accepted 16 January 2024
Available online 20 January 2024

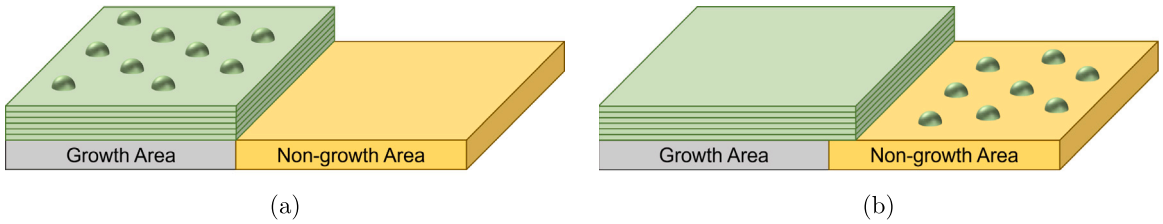

Fig. 1. Nonideal atomic layer deposition generating (a) nonuniform surface uniformity in the growth area and (b) growth in the non-growth area of the substrate.

growth on non-growth areas, which are sections of the substrate that do not require deposition, as depicted in Fig. 1. Thus, atomic layer etching (ALE) processes were developed to improve the film surface quality in a top-down fabrication approach, but results in additional processing time, which reduces the overall product throughput. A manufacturing solution, known as area-selective atomic layer deposition (AS-ALD), is an alternative procedure that deposits precursor reagents that only bind to the growth areas of the substrate due to the addition of a chemoselective inhibition step that hinders precursor adsorption on non-growth areas (Oh et al., 2021). AS-ALD is desirable in industrial contexts due to its creation of self-aligned structures that facilitate transistor stacking and densification (Kim et al., 2023). To ensure the efficacy of AS-ALD, and to enable this process in industrial applications, optimal operating conditions must be determined through experimental data, which is time-consuming. Semiconductor wafer processing comprises around 500 steps (Richard, 2023) that introduce a burden for high-throughput generation due to the complexities and demand for each manufacturing process (Leslie, 2022), which is consequential for supply chain shortages in electronics industries. While atomic layer deposition is only one step in the fabrication of the wafer, this process is crucial to the development of more complex geometries as semiconductor devices continue to be downscaled (Cremers et al., 2019). To address the challenges of experimental modeling, the application of *in silico* modeling is beneficial for alleviating the cost burden imposed by the reagents used for thin-layer deposition processes (Chien et al., 2022). Thus, this work proposes an *in silico* multiscale modeling framework to sufficiently characterize the AS-ALD of an Al_2O_3/SiO_2 (non-growth/growth area) substrate, and design a reactor configuration to allow for the scale-up of this process in relevant industrial settings.

The development of multiscale models for thin-layer deposition processes are beneficial to the generation of large datasets, but they require a complex cross-platform programming network that couples various simulations into a single framework (Maroudas, 2000). Multiscale models apply a combination of atomistic modeling through *ab initio* quantum mechanics computations to evaluate molecular and kinetic property data, mesoscopic modeling to characterize the stochastic surface kinetics through kinetic Monte Carlo methods, and macroscopic modeling to study the spatiotemporal behavior of fluids through computational fluid dynamics. This type of *in silico* modeling framework is beneficial towards studying the behavior of the AS-ALD process through various time and length scales and towards optimizing reactor configurations with large datasets. This work will study the effects of the reactor geometry by examining multiple discrete feed reactor configurations with the goal of determining the optimal delivery system to produce a high-quality thin film with minimal processing time.

This work is organized as follows: Section 2 examines the atomistic modeling of structural, electronic, and thermophysical properties and the mesoscopic modeling of surface scale kinetics, Section 3 discusses the development of the macroscopic CFD model of multiple AS-ALD reactor configurations through Ansys Fluent, Section 4 elucidates the multiscale modeling methodology used to conjoin the atomistic-mesoscopic and macroscopic simulations, and Section 5 analyzes the multiscale simulation results to determine the optimal reactor geometry that yields minimal process time for achieving full coverage and surface uniformity. A nomenclature of variables defined in this work is also provided in Table 1.

Table 1
Table of variables with their respective definitions and units.

Variable	Definition	Units
A_{site}	Surface area of active site	m^2
δt	kMC time advancement	s
Δt	CFD timestep size	s
E	Internal energy of the system	J
E_{act}	Activation energy	kJ mol^{-1}
$\vec{F}$	External body force	N m^{-2} s^{-1}
γ_1, γ_2	Uniform random number	–
$\vec{g}$	Gravitational acceleration of Earth	m s^{-2}
h	Planck's constant	J s
h_j	Sensible enthalpy	J kg^{-1}
$\vec{J}$	Diffusion flux rate	kg m^{-2} s^{-1}
k_B	Boltzmann constant	J K^{-1}
$k_{ads,s}$	Adsorption reaction rate constant	s^{-1}
k_{nonad}	Nonadsorption reaction rate constant	s^{-1}
k_{tot}	Sum of L possible reaction rate constants	s^{-1}
m_s	Atomic mass of gaseous reagent	g mol^{-1}
n	Number of active sites	–
P, P_s	Surface pressure for species s	Pa
ρ	Density of gas	kg m^{-3}
R	Ideal gas constant	J mol^{-1} K^{-1}
σ_s	Sticking coefficient for adsorbate s	–
S_h	Source energy flux generation rate	J kg m^{-3} s^{-1}
S_m	Source mass flux generation rate	kg m^{-3} s^{-1}
T	Absolute temperature of ambient environment	K
$\bar{\bar{\tau}}$	Stress tensor	N m^{-2}
$\vec{v}$	Fluid velocity	m s^{-1}
Z_s	Coordination number for gas s	–

2. Atomistic and mesoscopic modeling

A vital component to understanding the AS-ALD process lies in the kinetics of the surface reactions. In this work, a kinetic Monte Carlo model (kMC) is used to characterize the stochastic nature of surface reactions and determine their dependency on pressure and temperature. This procedure is conducted by first using atomistic modeling techniques via *ab initio* quantum mechanics simulations to derive the reaction rates of possible surface reactions. Then, a kMC algorithm is developed that replicates surface kinetics through a user-defined grid that represents a larger swath of the wafer surface and determines probable reaction pathways for each site on the grid.

The mesoscopic model is one of two integral components in the overall multiscale simulation. Based on the partial pressures and temperature on the surface of the wafer, the extents of reaction in one integration timestep, 0.001 s, are simulated. From this information, the macroscopic model can calculate how much reagent is consumed and how much product is produced, which is accounted for in subsequent timesteps.

2.1. Reaction rate calculations

The AS-ALD process examined in this work comprises three steps: (A) inhibition, (B) precursor adsorption, and (C) oxidation cycle. This work studies the AS-ALD of an Al_2O_3/SiO_2 substrate using acetylacetone (Hacac) as a small molecule and gaseous inhibitor for Step A, bis(diethylamino)silane (BDEAS) as a gaseous precursor for Step B,

and ozone (O_3) as a gaseous oxidant for Step C. To characterize all three steps of the AS-ALD process, this work simplifies the complex reaction mechanisms by concentrating on rate-limiting reaction steps determined through *in silico* modeling works (Mameli et al., 2017; Merkx et al., 2020; Yun et al., 2022a) and considering each reaction step as an elementary reaction.

All the reactions involved in the AS-ALD process can be classified into two types: adsorption and nonadsorption. Adsorption reactions can be modeled as bimolecular reactions, and as a result, their reaction rate constants, $k_{ads,s}$, for an adsorbate s, can be calculated through Collision Theory of gases. The aforementioned pressure and temperature-dependent formulation is described as follows:

$$k_{ads,s} = \frac{P_s A_{site} \sigma_s}{Z_s \sqrt{2\pi m_s k_B T}} \tag{1}$$

where P_s is the partial pressure of the gaseous reagent s, A_{site} is the surface area of a single active site, σ_s is an experimentally determined sticking coefficient unique to the reagent s, Z_s is the coordination number of the gas s, m_s is the atomic mass of the gaseous reagent s, and k_B is the Boltzmann constant.

The reaction rate constants of the nonadsorption reactions, k_{nonad}, are calculated with the temperature-dependent Arrhenius equation, as defined by the following equation:

$$k_{nonad} = \nu \exp\left(-\frac{E_{act}}{RT}\right) \quad \text{where } \nu = \frac{k_B T}{h} \tag{2}$$

where h is the Planck constant, E_{act} is the activation energy of the reaction, R is the universal gas constant, and T is the absolute temperature of the reaction. The pre-exponential factor, ν, is calculated using Transition-State Theory (TST) by assuming that the ratio of the partition functions for the transition state and the reactants is unity (Jansen, 2012). This assumption was validated with experimentally determined process times for observing full surface coverage by Merkx et al. (2020). The activation energy is found by first using *ab initio* quantum mechanics computations to optimize molecular and crystalline structures via Density Functional Theory (DFT). Then, the activation energies between the reactants and products are determined through Nudged Elastic Band (NEB) calculations. The aforementioned computations were conducted through the open-source electronic-structure optimization software Quantum ESPRESSO (QE) in a previous work by Yun et al. (2022a). It is notable that all of these k values are the reaction rates for a single active site and have units of s^{-1}.

2.2. Surface kinetics modeling

The kMC algorithm is a stochastic method that uses a set of randomly generated numbers to simulate the random nature of mesoscopic surface reactions in a spatiotemporal manner (Cheimarios et al., 2021). The algorithm used in this work is based on the BKL formulation established by Bortz, Kalos, and Lebowitz (Bortz et al., 1975) and has been modified to provide additional insight into the process when the process has larger reaction times relative to the timestep of the overall multiscale simulation, which is elucidated later on in this section. Prior to the execution of the BKL method in the kMC script, pressure and temperature data extracted from the computational fluid dynamics (CFD) simulation is utilized for the calculation of the reaction rate constants $k_{ads,s}$ and k_{nonad}. Next, the BKL method is conducted by assuming that potential reaction pathways for a given configuration are characterized by a Poisson distribution such that the reaction rate constants, k_i, from a list of possible reactions $i = 1, 2, \ldots, L$ are summed to enable the random selection of reaction pathways. This selection is made possible through the use of a uniformly calculated random number. Subsequently, a time advancement, i.e., the time in which the configuration transitions to the next state, is evaluated using a secondary uniform random number, which is summarized in Algorithms 1 and 2.

The BKL implementation in prior work (Yun et al., 2023) was conducted in the Python programming language using a grid-dependent procedure where the formulation, which is summarized in Algorithm 1, performs an iterative and ordered BKL procedure for each element in an $X \times Y$ grid. The result after each kMC execution produces grid data where the configuration or state of each element of the grid is discernible. However, this approach consumes memory that is not possible for implementation in user-defined functions (UDFs) in the computational fluid dynamics software, Ansys Fluent, and the integration of Python through cross-platform programming scripts reduces simulation efficiency.

To resolve the aforementioned challenges, a modification to Algorithm 1 is described in Algorithm 2 that contains the critical steps of the kMC methodology employed in this work, which is performed with C programming macros in UDFs that allow the program to communicate with the macroscopic simulation. Specifically, there are functions that use the macroscopic pressure and temperature data to calculate the reaction rates, which are used as inputs for the kMC algorithm. In an effort to preserve memory, a distinguishing feature of Algorithm 2 is the use of a random site selection method where site data is stored in variables that are independent of the location, i.e., the index, in the grid. A more detailed discussion of the memory constraints in UDFs is elucidated later in this section. Additionally, after the algorithm is completed for a given timestep, the new grid data is converted into the change in coverage, which is then used to calculate the generation and consumption source terms for the reactants and products.

There are two major concerns that the modified kMC algorithm must address. The first appears when one of the reactions has a k value magnitude that is smaller than that of any other reaction; i.e., the rate-limiting step is substantially slower compared to the remaining reactions. A consequence of this order-of-magnitude difference is that reactions that are more *rate-determining* will contribute more to the overall processing time. For instance, in Line 19 of Algorithm 1, which is from the algorithm of the BKL kMC method, the size of the time advancement is directly proportional to $1/k$. Thus, when $1/k$ is larger than Δt, the large time advancements will generate incomplete data in the form of a step-wise appearance because the algorithm will concurrently advance major portions of the grid concurrently whenever the rate-limiting step is the only available step.

The second area of concern is the memory usage. The BKL method employs a spatiotemporal approach to study the evolution and conversion of active sites that characterize the surface morphology of the substrate. For example, DeVita et al. (2005) simulated epitaxial growth using a kMC grid to study the surface morphology after each epitaxial cycle but observed computational constraints that limited the grid sizing to an $X \times Y$ grid. In a prior work (Yun et al., 2023; Tom et al., 2023b), the kMC algorithm was implemented through an external Python program and conjoined with the macroscopic computational fluid dynamics simulation in Ansys Fluent through a Linux Bash script.

In this work, the kMC algorithm was directly implemented in Fluent through custom User-defined Functions (UDFs) written in the C programming language (ANSYS, 2022c). UDFs are constructed by integrating Ansys-specific macros that allow extraction and manipulation of nodal data by defining integers attributed to boundaries in the simulation file (ANSYS, 2022a). UDFs also enable the declaration of user-specified variables, referred to as User-defined Memory (UDM) variables, with some restrictions that are detailed as follows. Specifically, following the aforementioned procedures, various macros are employed to enable parallelized computation, which is related to the partitioning procedure employed by the CFD software, iterative computation that is conjoined to the CFD solver, initialization, and executable actions, where C-based programming language is invoked for customized programs for the kMC code. Following the creation of the UDF-script written as a C program, the entire program is compiled through Ansys Fluent to generate the executable C programs, which work in tandem with the CFD simulation.

Algorithm 1: Original BKL kMC algorithm in Python with ordered site selection procedure.

Parameters: $P(x, y, t)$, $T(x, y, t)$ ▷ *Determined from CFD*

Input: Grid data, $k_{ads,s}(P, T)$, $k_{nonad}(T)$ ▷ *Calculated from Parameters*

Output: Grid data, δt

```
▷ Let there be X grid rows and Y grid columns
▷ Let there be L reactions in the process, and let k_i represent the ith reaction
δt = 0
while δt < Δt do                                   ▷ Running algorithm until kMC timestep is as large as CFD timestep
    for each species do
        if number of species in grid = 0 then
            Set the appropriate k_i value(s) to 0   ▷ Removing impossible reactions
    k_tot = Σ_{i=1}^{L} k_i
    for j in 1 : X do
        for k in 1 : Y do
            ▷ Randomly determining if a reaction occurs for each site on the X × Y grid
            Randomly select γ_1 ∈ (0, 1]
            for r in 1 : L do
                if Σ_{i=1}^{r-1} k_i < γ_1 k_tot ≤ Σ_{i=1}^{r} k_i then
                    if reaction r is possible then
                        Execute reaction r
                        ▷ Steric hindrance for Step B is not shown here.
    Randomly select γ_2 ∈ (0, 1]
    δt = δt − ln(γ_2)/k_tot                        ▷ Advancing kMC timestep
```

Algorithm 2: Modified kMC algorithm in UDF with random site selection procedure.

Parameters: $P(x, y, t)$, $T(x, y, t)$ ▷ *Determined from CFD*

Input: Grid data, $k_{ads,s}(P, T)$, $k_{nonad}(T)$ ▷ *Calculated from Parameters*

Output: Grid data, δt

```
δt = 0
while δt < Δt do                                   ▷ Running algorithm until kMC timestep is as large as CFD timestep
    Randomly select a site on the grid
    ▷ Let there be L possible reactions for the selected site and let k_i represent the ith reaction
    k_tot = Σ_{i=1}^{L} k_i
    Randomly select γ_1, γ_2 ∈ (0, 1]
    for r in 1 : L do                              ▷ Going through each possible reaction to randomly select one
        if Σ_{i=1}^{r-1} k_i < γ_1 k_tot ≤ Σ_{i=1}^{r} k_i then
            Execute reaction r
            ▷ Steric hindrance for Step B is not shown here
    Determine n, the number of active sites from grid data
    δt = δt − ln(γ_2)/(n k_tot)                    ▷ Advancing kMC timestep
```

While this approach provides a major increase in computational performance, this method also comes with more restrictions for the new code. The main restriction when implementing UDFs in Fluent is the memory storage, where a maximum of 500 variables can be safely stored for each integration timestep (ANSYS, 2022c). However, each kMC grid is defined by a 300 × 300 lattice in this work for a total of 90,000 sites. Thus, the data of each kMC grid must be stored differently so that it can be represented by less than 500 variables.

To rectify the two issues presented, the kMC algorithm used in this work was modified to evaluate the time required for a single active site to react, whereas the BKL algorithm evaluates how the entire grid progresses in a given timeframe (Bortz et al., 1975). This difference is implemented in the form of three distinctive changes, which are summarized as follows:

(1) The arrangement of occupied sites is neglected by employing a Markov chain in which only one site on the entire grid advances with each step as conducted by Kim et al. (2020).
(2) The time advancement computation includes the number of unoccupied sites on the grid, where a reduction in unoccupied sites increases the time progression as employed by Giménez et al. (2002) and Klement et al. (2021).
(3) Grid data is stored as a single integer variable that counts the number of each species, rather than as an array.

The first two modifications resolve the first concern regarding small k values, and the last modification resolves the second concern regarding memory storage. These adjustments to the BKL formulation are further examined and verified in Section 2.2.1.

2.2.1. Derivation of the modified kMC algorithm

To properly make the necessary modifications to the BKL kMC algorithm, it is first important to understand what the kinetic rate represents. Intuitively, the parameter reflects the number of reactions that occurs every second at an active site, which is a region on the substrate surface that is able to undergo a chemical reaction. In other words, the kinetic rate is a measure of the probability that such an event is happening. Additionally, these events can be assumed to follow a Poisson distribution, to react independently of other sites, and to react independently of other possible reactions (Christofides and Armaou, 2006). As such, probability theory allows the reaction rate to be decomposed into two independent events (Gillespie, 1976), as shown in Eq. (3). First, there is the probability that the site is in a state where the reaction can proceed, $\mathcal{P}(possible)$. Second, there is the probability that the reaction actually proceeds, $\mathcal{P}(proceed)$. This decomposition is expressed as follows:

$$k_{rxn} = \mathcal{P}(reaction)$$

$$k_{rxn} = \mathcal{P}(possible) \cdot \mathcal{P}(proceed) \tag{3}$$

where k_{rxn} is the reaction rate of a given reaction for a single active site as calculated in Section 2.1 and $\mathcal{P}(reaction)$ is the probability of that reaction taking place in one second.

A similar expression for the reaction rate of the entire grid can be derived by using the number of sites rather than the probability of a site being able to undergo the desired reaction. This expansion is also based on the assumption that these two events are mutually exclusive, yielding:

$$k_{grid} = \mathcal{N}(possible) \cdot \mathcal{P}(proceed) \tag{4}$$

where k_{grid} is the average reaction rate for a given kMC grid and $\mathcal{N}(possible)$ is the number of sites that can undergo the desired reaction.

Because of the assumption that each site on the grid is independent of the others, $\mathcal{P}(possible)$ can be related to $\mathcal{N}(possible)$ as follows:

$$\mathcal{N}(possible) = n \cdot \mathcal{P}(possible) \tag{5}$$

where n is the number of active sites on the grid. With this relationship, k_{grid} can be related to k_{rxn} as follows:

$$k_{grid} = \mathcal{N}(possible) \cdot \mathcal{P}(proceed)$$

$$k_{grid} = n \cdot \mathcal{P}(possible) \cdot \mathcal{P}(proceed)$$

$$k_{grid} = n \cdot k_{rxn} \tag{6}$$

where k_{grid} is the average reaction rate of a given reaction for the entire grid, k_{rxn} is the reaction rate for a single active site, and n is the number of actives sites on the grid.

When there are multiple possible reactions, it is necessary to determine k_{tot}, which is the probability that each reaction is independent of that of other reactions. Since it is assumed that the probability of each k is independent of other reactions, the probability that one of two reactions will take place can be found using the inclusion-exclusion principle as follows:

$$\mathcal{P}(k_i \cup k_j) = \mathcal{P}(k_i) + \mathcal{P}(k_j) - \mathcal{P}(k_i \cap k_j)$$

where $\mathcal{P}(k_i \cup k_j)$ is the probability that either reaction i or reaction j will take place, $\mathcal{P}(k_i)$ is the probability that reaction i will occur, $\mathcal{P}(k_j)$ is the probability that reaction j will occur, and $\mathcal{P}(k_i \cap k_j)$ is the probability that both reaction i and reaction j will occur. This equation can be simplified by noting that $\mathcal{P}_i = k_i$ as shown in Eq. (3) and that $\mathcal{P}(k_i \cap k_j) = 0$ because the two reactions are mutually exclusive, which yields:

$$\mathcal{P}(k_i \cup k_j) = k_i + k_j$$

If there are L total possible reactions, these additional reactions can be summed into $\mathcal{P}(k_i \cup k_j)$ to obtain the probability that an unspecified reaction occurs as all of these reactions are independent and mutually exclusive. The expression for this possibility is:

$$\mathcal{P}(k_{tot}) = k_{tot} = \sum_{z=1}^{L} k_z \tag{7}$$

where k_{tot} is the possibility of an unspecified reaction occurring and k_z represents the possibility for a specific reaction to occur. Note that, when evaluating an active site in isolation, $k_z = k_{rxn}$; similarly, when evaluating an active site in the context of the entire grid, i.e., $k_z = k_{grid}$. Thus, to evaluate the time progression for a single active site at a time, $k_{tot,grid}$ can be represented as follows:

$$\begin{aligned} k_{tot,grid} &= \sum_{z=1}^{L} k_{grid} \\ &= \sum_{z=1}^{L} n \cdot k_{rxn} \\ &= n \cdot \sum_{z=1}^{L} k_{rxn} \\ &= n \cdot k_{tot,rxn} \end{aligned}$$

where $k_{tot,grid}$ is the possibility that an unspecified reaction will occur anywhere on the kMC grid, $k_{tot,rxn}$ is the possibility that an unspecified reaction will occur at a singular active site, and n is the number of active sites on the kMC grid. This formula is employed in Line 12 of Algorithm 2.

The other modification made to the kMC method presented in Yun et al. (2022a) is the reduction of memory usage as necessitated by the restrictions of Ansys Fluent. The total data stored in between each timestep must be reduced from 90,000 integers for a 300 × 300 grid to less than 500 integers. This procedure was accomplished by taking advantage of the fact that the kMC algorithm does not use any positional data; i.e., the simulation is not concerned about the location of the site in the grid, but rather only the state of the grid for each timestep. Thus, instead of using an array with 90,000 entries, the amount of

Algorithm 3: Step B steric hindrance locking algorithm.

Variables:

V4: Completely reacted site

IC: Site that will never reach completion

LK: 2 adjacent V4s that cannot trap an unfinished site

▷ *Let there be N total sites*

▷ *Let S_i represent the species of site i*

for j in 1 : N **do** ▷ *Going through each site on the grid*

  Randomly select a species S_j

  ▷ *Let S_{adj1}, S_{adj2} be two randomly selected sites that represent the sites adjacent to S_j*

  if $S_j \neq$ V4 AND S_{adj1} = V4 AND S_{adj2} = V4 **then**

    $S_j \longrightarrow$ IC

  else if S_j = V4 AND (S_{adj1} = V4 OR S_{adj2} = V4) **then**

    $S_j \longrightarrow$ LK

    $V4 \longrightarrow$ LK ▷ *This represents the adjacent V4 turning into LK*

each intermediate species was counted and saved in its own variable. For example, in Step C of the AS-ALD cycle, there are 3 species: V4, V6, and V8. The old method would have 90,000 entries, each one representing a site and tracking whether it is in state V4, V6, or V8. The new method has 3 variables, which are defined as *buckets* that represent the number of V4 sites, the number of V6 sites, and the number of V8 sites. By discarding the unnecessary positional data of the sites, this bucket method is able to store the relevant information of all the kMC grids in 31 variables where 25 variables are reserved for tallying species involved in the Steps A, B, and C of the AS-ALD process and the remaining 6 variables are used for defining the source generation and consumption flux rate terms on the wafer surface boundary conditions.

However, the reaction mechanism for Step B is more complicated than the other two reactions, as steric hindrance plays an essential role in the kinetics of that process. The specific details of the effects of steric hindrance on the BKL formulation can be found in Yun et al. (2022a), but a summary is as follows. Each site has two adjacent neighbors that apply two conditions on the surface reaction mechanism. The first extra condition is that a site is restricted from certain reactions if a bulky molecule has adsorbed to either neighboring site. Physically, these molecules hinder the primary site from reacting. The second condition is attributed to the final surface reaction, which requires 2 adjacent sites to bond and for both sites to reach the final state. Thus, it is possible for situations to arise where it is impossible for a site to fully react if both of its neighbors have reached completion by bonding with other sites. As a result, these sites must be deactivated so that the kMC algorithm can reach completion.

To implement the first condition described above, the modified kMC algorithm creates buckets that represent the number of adjacent sites that are blocking the site of adsorption. Because each site has two sterically relevant neighboring sites, there are 3 block status buckets: unblocked, one-block, and two-block. While both the one-block and two-block statuses represent the target site being unable to undergo certain surface reactions, distinguishing between the two allows the preservation of more positional data about the grid. During the kMC algorithm, whenever the algorithm needs to determine whether the site it randomly selected is blocked, it will do so by randomly selecting a block status. The second condition is an extra step that takes place after each kMC event. To properly deactivate sites, the procedure described in Algorithm 3 is executed after each iteration of Algorithm 2.

By running Algorithm 3, the number of trapped sites that are unable to reach the final V4 state can be accurately represented even after discarding all positional data. After implementing both modifications discussed in this section, the kMC model is able to simulate the surface reactions with greater resolution and obtain high quality results for all the reactions in the AS-ALD process.

To verify that the results of the modified kMC algorithm are accurate and valid, comparisons between Algorithms 1 and 2 and were examined for both cases with slow reactions (Step B) and cases without slow reactions (Steps A and C) in Fig. 2.

3. Computational fluid dynamics modeling

Computational fluid dynamics (CFD) simulations describe the macroscopic behavior of fluids in larger time and length scales, which enables the scale-up of processes. The integration of CFD is applicable to characterizing the spatiotemporal flow of reagents on the substrate surface, which experiences surface reactions that consume the reagents and generate byproducts. The development of a CFD model requires the construction of a computer-aided design (CAD) model for a three-dimensional (3D) discrete feed reactor system, the performing of a meshing procedure on the CAD model, and the creation of the CFD simulation of the AS-ALD process in the discrete feed reactor model.

3.1. Reactor design

Several reactor models have been investigated to discuss the uniformity of reagent coverage on the substrate surfaces and to improve the productivity of the process. For example, Yun et al. (2022b) proposed a cross-flow reactor to control the behavior of flow in the azimuthal direction of the substrate for an atomic layer etching process. Additionally, Dobkin and Zuraw (2003) and Elers et al. (2006) suggested using showerhead reactors to improve the uniformity of fluid flow in the radial direction. By considering the challenges attributed to low product throughput, Yun et al. (2023) proposed a spatial reactor configuration where the reagent is delivered perpendicularly to the substrate in a continuous feeding mechanism for atomic layer etching and area-selective atomic layer deposition processes. While the aforementioned reactor models have effectively yielded valuable results in improving product

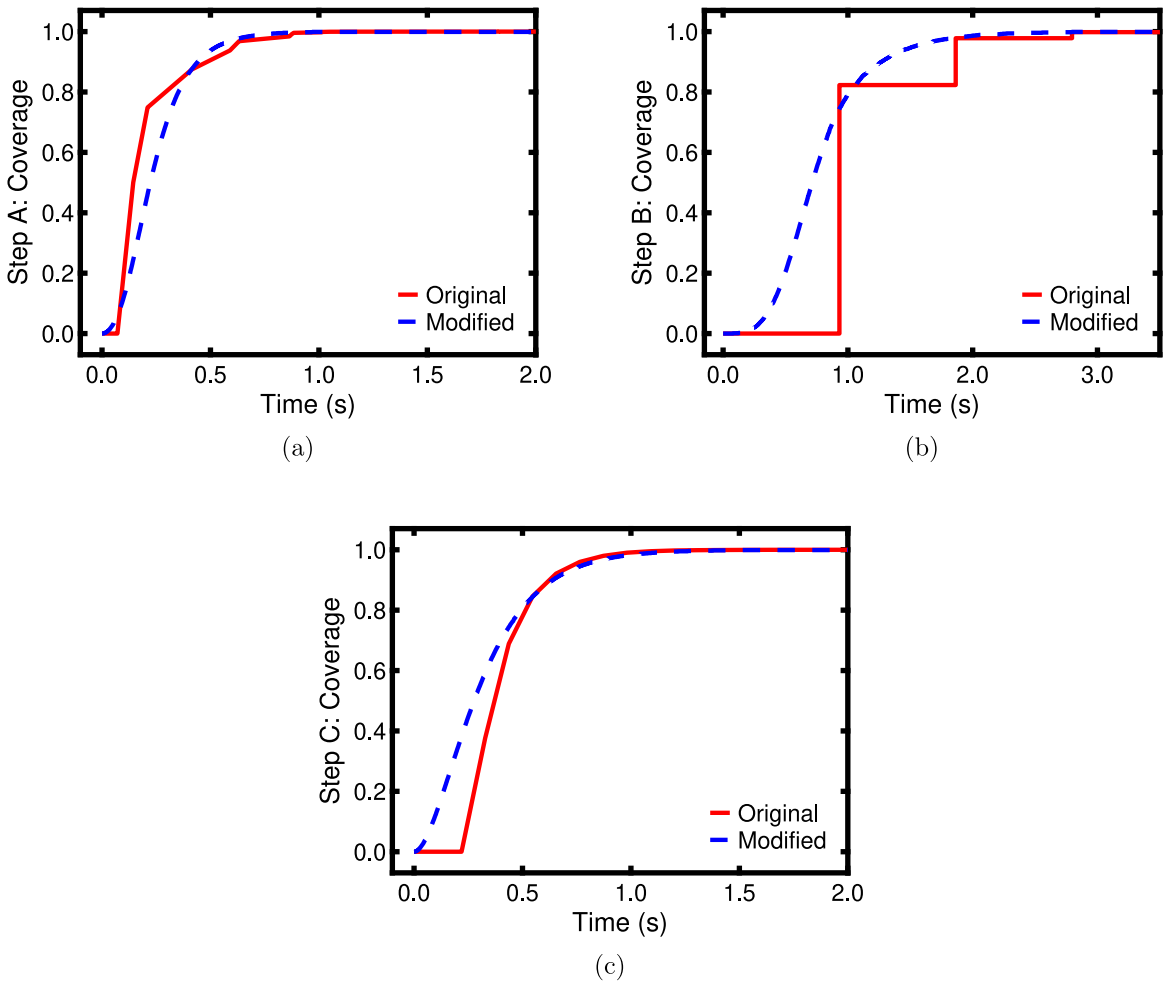

Fig. 2. Comparison of the original (Algorithm 1) and modified (Algorithm 2) kMC algorithms for Steps (a) A, (b) B, and (c) C. The original kMC processing times to reach full coverage for Steps A, B, and C are 1.018, 2.796, and 1.418 s, respectively. The modified kMC processing times to reach full coverage for Steps A, B, and C are 0.969, 2.793, and 1.487 s, respectively.

quality and yield, this work considers the impact of steric collisions generated from bulky molecular species including Hacac and BDEAS, which introduces challenges associated with surface uniformity. Thus, there is motivation to develop a reactor that minimizes steric hindrance induced by screening effects.

This work adopts a previously designed discrete feed reactor (Tom et al., 2023a) inspired by the work of Lin et al. (2023) through Ansys DesignModeler, which delivers reagent perpendicularly to the substrate surface in discrete pulses through an injection plate. The employment of discrete feeding with cut-in purging allows the byproduct species that inhibit adsorption of Hacac and BDEAS on the substrate surface to be regularly removed. The discrete feed reactor, illustrated in Fig. 3, situates a showerhead divider that is below and parallel to the injection plate to facilitate the transport of reagents in the radial directions of the substrate, thereby maximizing the exposure of the substrate to the reagent in minimal pulse times. The gap distance between the injection to the showerhead plate is 3 mm, and the gap distance between the showerhead plate to the 200-mm diameter substrate surface is 5 mm. These gap distances are necessary to minimize the volume required to maintain laminar flow behavior (Kim et al., 2003). A summary of the reactor dimensions are presented in Table 2. The injection plate has a substantial impact on the mass transport of reagents in the radial direction. Thus, various injection plate geometries, which are illustrated in Fig. 4, were previously proposed by Tom et al. (2023a) to observe their impact on the fluid dynamics on the substrate surface. Results from the aforementioned work provided valuable information about the role of characteristic lengths on the rate of mass transfer in the radial direction. This work extends prior macroscopic modeling work for each reactor injection plate geometries by studying their effect on the spatiotemporal coverage and process time required to reach complete surface coverage (see Table 3).

Table 2
Dimensions for the reactor configurations[a].

Reactor dimension	Value
Plate Diameter	290 mm
Ring Inlet Outer Diameter	170 mm
Ring Inlet Inner Diameter	130 mm
Round Inlet Diameter	20 mm
Round Outlet Diameter	4 mm
Showerhead Diameter	250 mm
Showerhead Pores Diameter	10 mm
Showerhead Plate Thickness	0.5 mm
Showerhead Plate/Wafer Vertical Gap Distance	5 mm
Reagent Inlet Plate/Showerhead Plate Vertical Gap Distance	3 mm
Wall Sector Angle	40°

[a] All dimensions are fixed for each reactor configuration.

3.2. Meshing

Following the construction of the reactor model, a discretization process is conducted to produce conformal meshes that balance

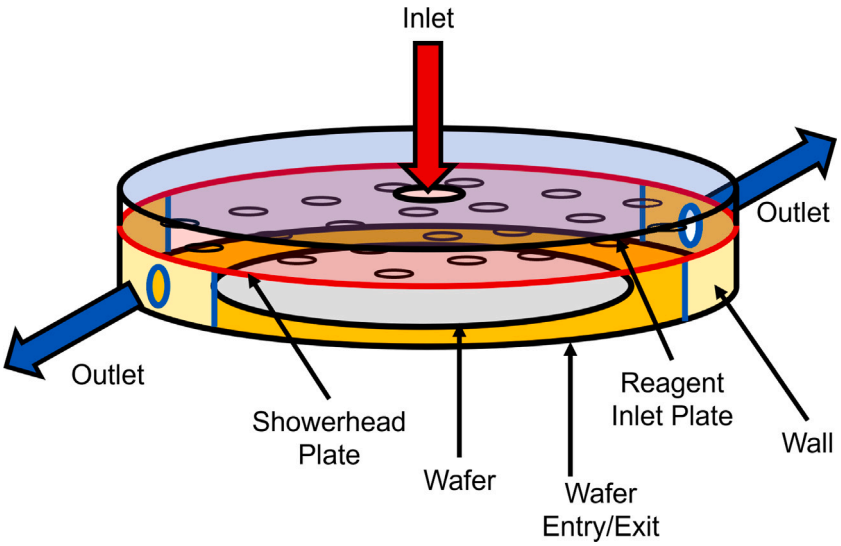

Fig. 3. Schematic of the discrete feed reactor model for the AS-ALD reaction. The operation of the reactor is conducted in sequential pulses where a mixture of Hacac and N_2, BDEAS and N_2, and O_3 and N_2 are injected in the inlet stream for Steps A, B, and C, respectively. Purged materials including Hacac, H_2, and N_2 for Step A, BDEAS, N_2, and DEA for Step B, and O, O_2, O_3, and N_2 for Step C, are evacuated through the outlets, which include the wafer exit/entry.

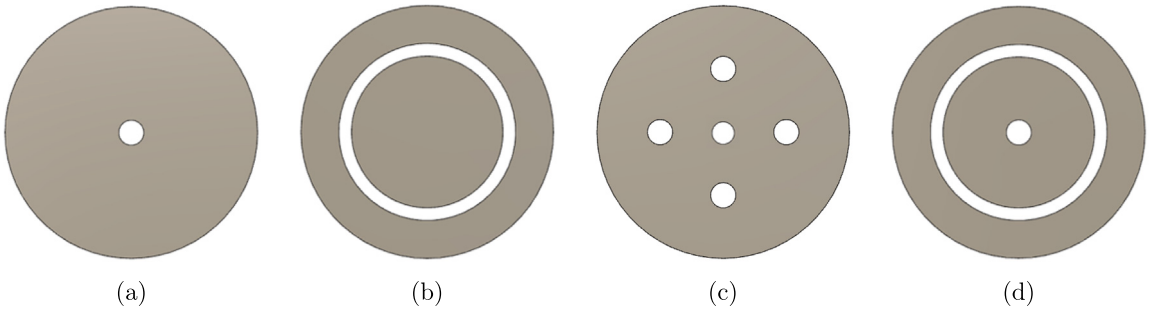

Fig. 4. Various feed distributor geometries for (a) Single, (b) Ring, (c) Multi, and (d) Combined reactor configurations.

computational efficiency and accuracy when performing the finite element method. Meshes for each reactor model are produced from "Meshing Mode", a feature of the multiphysics software, Ansys Fluent, in a prior work (Tom et al., 2023a). The aforementioned meshes were generated by optimizing mesh quality parameters based on the tetrahedral geometries of the discretized cells, which include the orthogonality, aspect ratio, and skewness (ANSYS, 2022b). To maximize each reactor configuration mesh, optional remeshing tools were then applied to the irregular surface and volume cells. In additional to maintaining balanced mesh quality, this work aims to minimize the number of cells required to produce the 3-D meshes to reduce the complexity of the computational fluid dynamics simulation, where each reactor configuration comprises 1.1 to 1.2 million cells.

3.3. Computational fluid dynamics simulation framework

The macroscopic CFD simulation is constructed by defining boundary, operating, and solver conditions that are specific to the AS-ALD process. This simulation will employ a strategy for solving the mass, momentum, and energy transport equations, which are described as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \vec{v}\right) = S_m \tag{8}$$

$$\frac{\partial}{\partial t}\left(\rho \vec{v}\right) + \nabla \cdot \left(\rho \vec{v}\vec{v}\right) = -\nabla P + \nabla \cdot \left(\bar{\bar{\tau}}\right) + \rho \vec{g} + \vec{F} \tag{9}$$

$$\frac{\partial}{\partial t}\left(\rho E\right) + \nabla \left(\vec{v}\left(\rho E + P\right)\right) = -\nabla \left(\Sigma h_j \vec{J}_j\right) + S_h \tag{10}$$

where the mass transport equation in Eq. (8) is associated with the gas-phase species flux, which is represented by the product of the gas-phase species density, ρ, and the velocity of the species, $\vec{v}$, and is related to the species source generation and consumption flux rate, S_m. The momentum transport equation in Eq. (9) relates the rate of momentum per unit volume to the convection, pressure, viscous, and gravitational forces where P is the operating pressure of the reactor, $\bar{\bar{\tau}}$ is the normal two-rank stress tensor, $\vec{g}$ is the gravitational acceleration constant, and $\vec{F}$ is the force acting on the system. The energy transport equation defined in Eq. (10) describes the relation of the accumulated rate of system energy, E, with the convective, conductive, and energy source generation or consumption, S_h, rates, where h_j and $\vec{J}_j$ is the sensible enthalpy and mass diffusion flux, respectively, of the gas species j.

Ansys Fluent contains multiple fluid dynamics models that can be used to describe the behavior of the fluid flow. Due to the small reactor sizes and observance of laminar behavior from prior research (Tom et al., 2023a), a laminar fluid model is defined in the simulation. The mass transport is simulated by specifying gas-phase reagent and byproduct species that are present in the Ansys ChemKin database and thermophysical property data generated from experimental works and *ab initio* quantum mechanics calculations discussed in Section 2. The source generation and consumption flux rate terms are evaluated through the kMC simulation and defined as boundary conditions on the wafer surface through user-defined functions (UDFs). Additionally, this simulation considered the role of ozone decomposition within the reactor and surface of the wafer. The reactor is also operated under isothermal and isobaric conditions by assuming that a temperature control system is used to maintain the temperature on the wafer surface

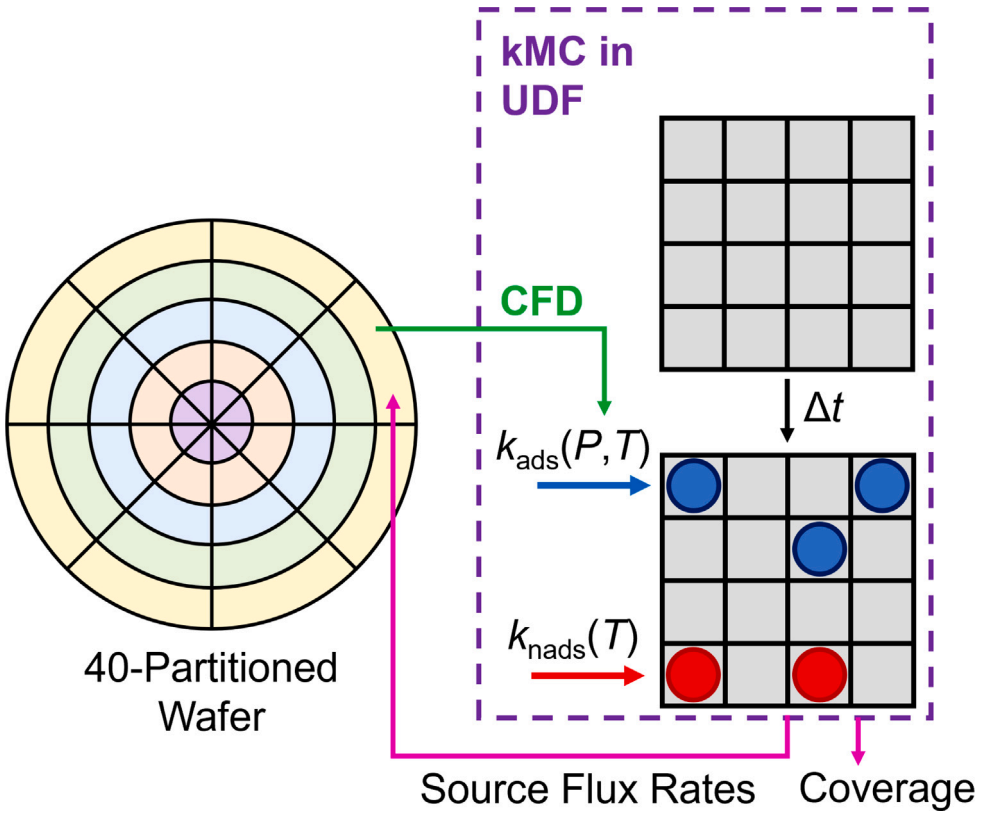

Fig. 5. Illustration of the multiscale CFD modeling framework. The wafer is partitioned into 40 sections in the CFD simulation to produce a collection of 40 surface pressure and temperature datasets that are used to calculate the reaction rate constants in user-defined functions (UDFs). The kMC simulation, which is performed in the UDF calculates the surface coverage and source flux rate terms that are transmitted to the CFD simulation.

Table 3
Operating conditions for each reactor geometry.

Reactor	Temperature (K)	Pressure (Pa)	Mole fraction			Mass flow rate (kg/s)
			Hacac	BDEAS	Ozone	
Single	573	300	0.50	0.50	0.20	2.00×10^{-5}
Ring	573	300	0.50	0.50	0.20	2.00×10^{-5}
Multi	573	300	0.50	0.50	0.20	Each Inlet: 4.00×10^{-6}
Combined	573	300	0.50	0.50	0.20	Single: 1.00×10^{-5} Ring: 1.00×10^{-5}

and that a vacuum pump is effectively applied to regulate the pressure within the reactor chamber.

A pressure-based coupled solver method is integrated into this work to simultaneously solve the momentum and pressure-based continuity equations in a parallelized algorithm to reduce computation time at a cost of increased memory requirement. To circumvent this issue, CPU-based (central processing unit) nodes were integrated into this work comprising 48 and 36 cores with 512 GB and 384 GB of dynamic random-access memory (DRAM), respectively, and executed through text-user interface (TUI) commands to minimize graphical power. Additionally, a fixed timestep method of step size $\Delta t = 0.001$ s is defined, which is within the Courant number threshold recommended by the default settings for the program. Lastly, under-relaxation factors of 0.5 were assigned to all gas-phase species involved in the mass transport calculations to minimize the potential for divergent or oscillatory residual responses that could potentially be generated from the source flux rate terms evaluated from the kMC simulation.

4. Multiscale modeling

The efficacy and impact that simulations can have naturally depends on their accuracy and precision. Generally speaking, the accuracy of a simulation can always be improved by increasing the computational costs; for example, lowering the integration timestep when numerically solving a differential equation will improve the accuracy of the final answer while increasing the number of calculations that must be made to reach that final answer. Thus, one of the driving motivations for this paper is finding an optimal balance between the accuracy and the computational cost of the simulation.

One commonly used method to improve simulation accuracy without an expensive computational cost is multiscale simulation (Wehinger et al., 2022). This method comprises two simulations that run concurrently and interact with each other: a mesoscopic kMC model that simulates the surface kinetics of the wafer as a function of the pressure and temperature, and a macroscopic CFD model that simulates how the pressure fields within a reactor evolve with time. These two interacting models improve the overall accuracy of the simulation because their domains are intrinsically linked. The surface reactions on the wafer generate and consume products and reagents, which affects the overall pressure fields in the reactor, which affects the reaction rate on the wafer surface. Thus, to improve simulation accuracy, the two models are integrated together in a multiscale framework as shown in Fig. 5.

At each integration timestep, $\Delta t = 0.001$ s, the CFD model takes the generation and consumption terms calculated by the kMC model

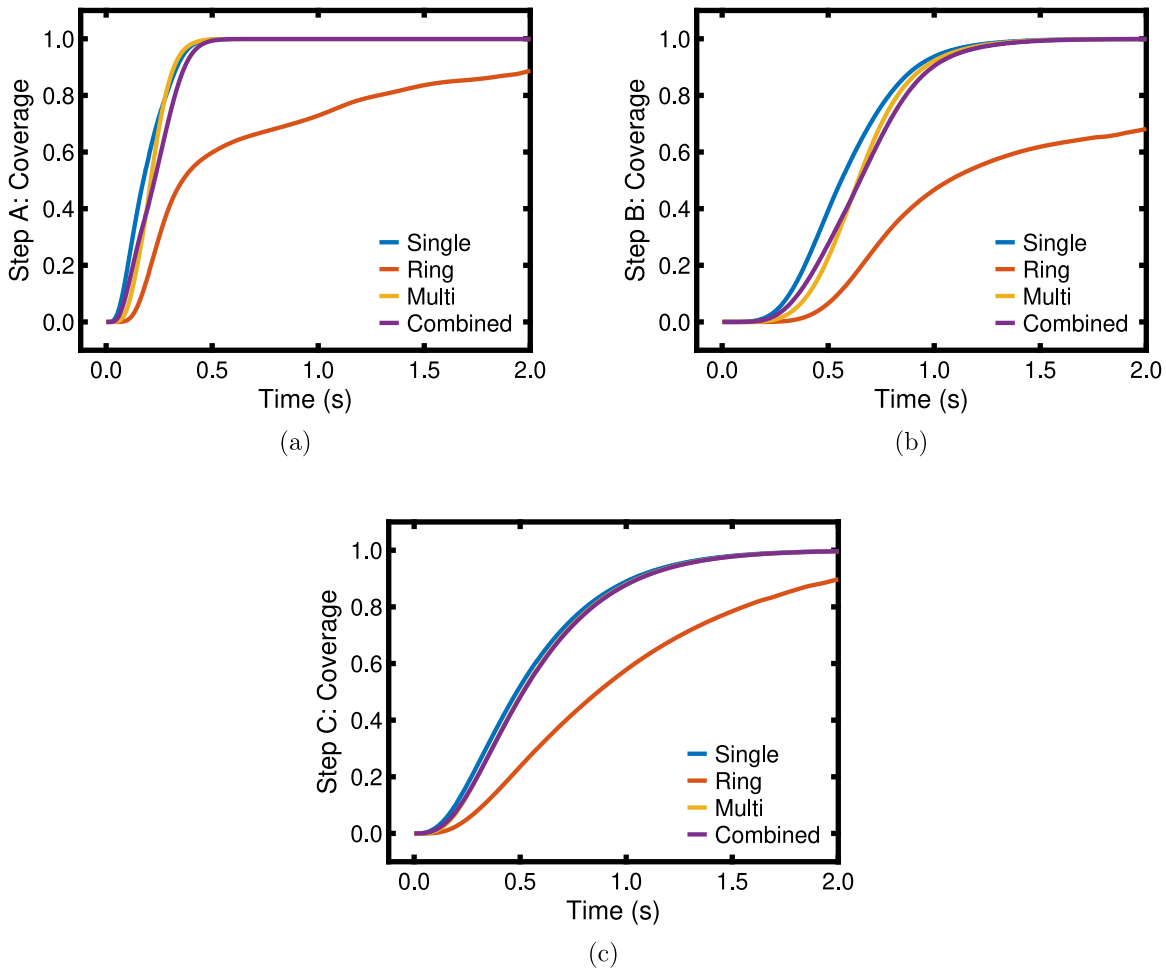

Fig. 6. Reactor configuration comparison of the temporal progression of the average surface coverage for (a) Step A, (b) Step B, and (c) Step C.

in the previous timestep into account when calculating the pressure fields in the reaction. Then, the kMC model receives information about the species pressure and temperature at the surface of the wafer and uses that to calculate the extent of any surface kinetics, as well as the resulting consumption and generation of species. After repeating this step for multiple timesteps, the multiscale simulation offers a comprehensive understanding of how the wafer surface reactions evolve as time progresses inside the reactor. With the computational resources and numerical simulation specifications stated above, the computing time ranged from 4 to 6 h for a process time of 3.5 s for the 48-core and 36-core nodes, which allows for effective data production.

5. Multiscale simulation results and discussion

The aim of the multiscale CFD simulation is to give a quantitative understanding of how various reactor designs affect the process efficiency of the AS-ALD reactions, which is directly tied to the process time. Due to the self-limiting nature of AS-ALD, the process is naturally resistant to over-deposition. Rather, the only drawbacks of overprocessing are unneeded consumption of the reagents and a decreased product throughput. Thus, there is a large economic incentive to minimize the process time, as this will both reduce the reagent consumption and increase production capabilities.

To evaluate the overall process efficiency, two main criteria are taken into account: the minimum process time and surface coverage uniformity. The minimum process time is the time at which the substrate achieves full coverage for the surface-terminated product, which is directly dependent on the reagent dosage time. On the other hand, the surface coverage uniformity is a measure of how the standard deviation of the coverage on the substrate surface changes with time. Naturally, the lower the overall standard deviation, the less reagent is wasted, and the reverse implies greater reagent wastage. These two criteria are positively correlated, as a low standard deviation distribution implies an effective usage of the reagent, which then implies that the process will be completed quickly.

Fig. 6 illustrates the temporal progression of coverage for Steps A, B, and C for each reactor configuration. Generally, the ring-shaped reactor geometry underperforms due to the vacuum pressure forces and steric screening effects from byproducts that result in a lack of fluid transport toward the center of the wafer. Meanwhile, the single and multi-shaped injection plates are conducive towards achieving full surface coverage in minimal processing times by concentrating the reagent toward the center of the wafer. With smaller characteristic lengths, especially for the combined reactor geometry, the injection flow rate is unable to overcome the effects of the vacuum pressure forces. Thus, an initial delay in coverage is observed for the coverage profiles of the combined model when compared to those of the single and multi feed reactors. Table 4 also summarizes the processing times required to achieve full surface coverage for each reactor configuration and each step in the AS-ALD ABC cycle, where the single and multi reactor models were observed to have the fastest processing times.

The reactor design with the smallest minimum process time can be quickly determined by examining the average coverage progression across the wafer as a function of time, as shown in Fig. 6. From Table 4, it can be seen that the multi-shaped reactor performed the best overall with a total minimum process time of 3.638 s. However, to understand

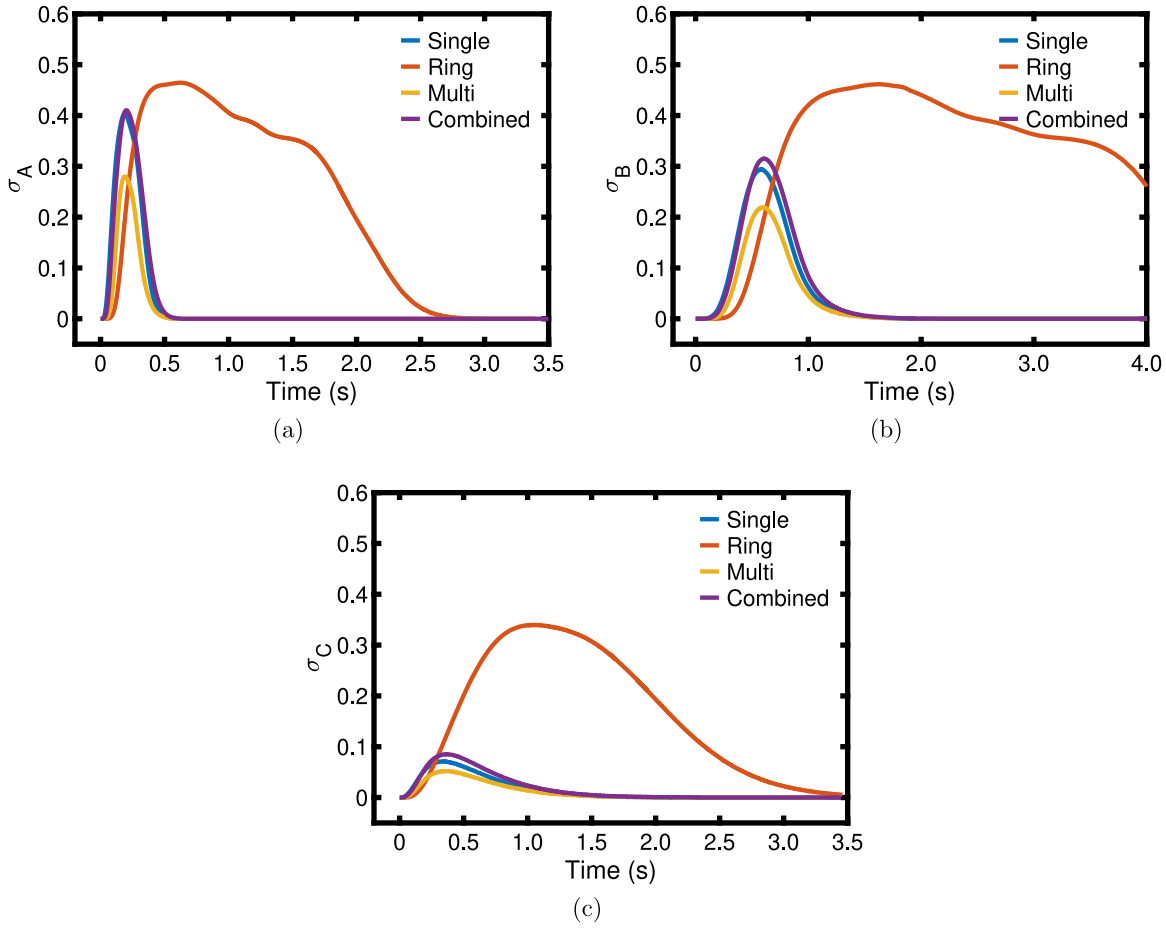

Fig. 7. Reactor configuration comparison of the temporal progression of the standard deviation, σ, in surface coverage for (a) Step A, (b) Step B, and (c) Step C.

Table 4
Computed process times required to obtain full surface coverage for the terminated products from Steps A, B, and C as a function of reactor configuration.

Reactor	Process time (s)			Total process time (s)
	Step A	Step B	Step C	
Single	0.466	1.431	1.783	3.680
Ring	2.310	4.586	2.596	9.492
Multi	0.453	1.448	1.737	3.638
Combined	0.509	1.508	1.755	3.772

why this reactor design performed the best, it is necessary to examine more specialized data, such as the standard deviation progression as a function of time and contour plots of the coverage at specific points in time.

The multi-shaped reactor, which has the smallest process times, also consistently has the lowest standard deviation for all points in time. However, the single and combined-shaped reactors have similar standard deviation curves, illustrated in Fig. 7, for all three reactions, but the single-shaped reactor consistently outperforms the combined-shaped reactor in terms of process time. The standard deviation plots demonstrate the uniformity of wafer surface coverage for Steps A, B, and C of the AS-ALD process, but lack critical information in determining the spatial dependence of the surface coverage on the injection geometry.

Meanwhile, the contour coverage plots directly reflect the progress of the surface reactions. From Eqs. (1) and (2), it can be seen that, in an isothermal reactor, the reaction rates are only a function of pressure. Thus, the progress of the surface reaction at a particular point is related to the pressure which that particular point has experienced since the start of the reaction. This suggests that the contour coverage plots effectively represent how ideal the pressure distribution of a given reactor design is. In Fig. 8, all the reactor designs have a high coverage value in the areas directly underneath the precursor dispensers. However, the most valuable information from these contour graphs are the areas with low coverages; these areas represent sections far from the inlet, where the reactor design plays a major role in how high the precursor pressure necessary for carrying out the reactions is. For example, the ring-shaped reactor design has a large area in the center with 0 coverage, which means that almost no precursor is flowing there. This matches the results in Tom et al. (2023a), which found that the pressure in the center is very low for the ring-shaped reactor. Meanwhile, out of the single, multi, and combined-shaped reactors, the multi-shaped reactor has the best coverage contour plot in that its outer edges have the highest coverage. Even though the coverage at the outer edge is the same for both the single and combined-shaped reactors, the coverage at the middle of the reactor, $r = (12, 16)$ cm, is higher for the single-shaped reactor. This means that the single-shaped reactor design has a better pressure distribution in the middle section of the wafer, rendering it superior to the combined-shaped reactor design. The results of the contour plots show that the multi-shaped reactor design is the best, followed by the single, combined, and the ring-shaped in that order. Similarly, the contour plots for Steps B and C, as shown in Figs. 9 and 10, demonstrate that the multi-shaped reactor performed the best with the other three reactor designs following in the order described above. This correlates with the performances noted in Table 4, which shows

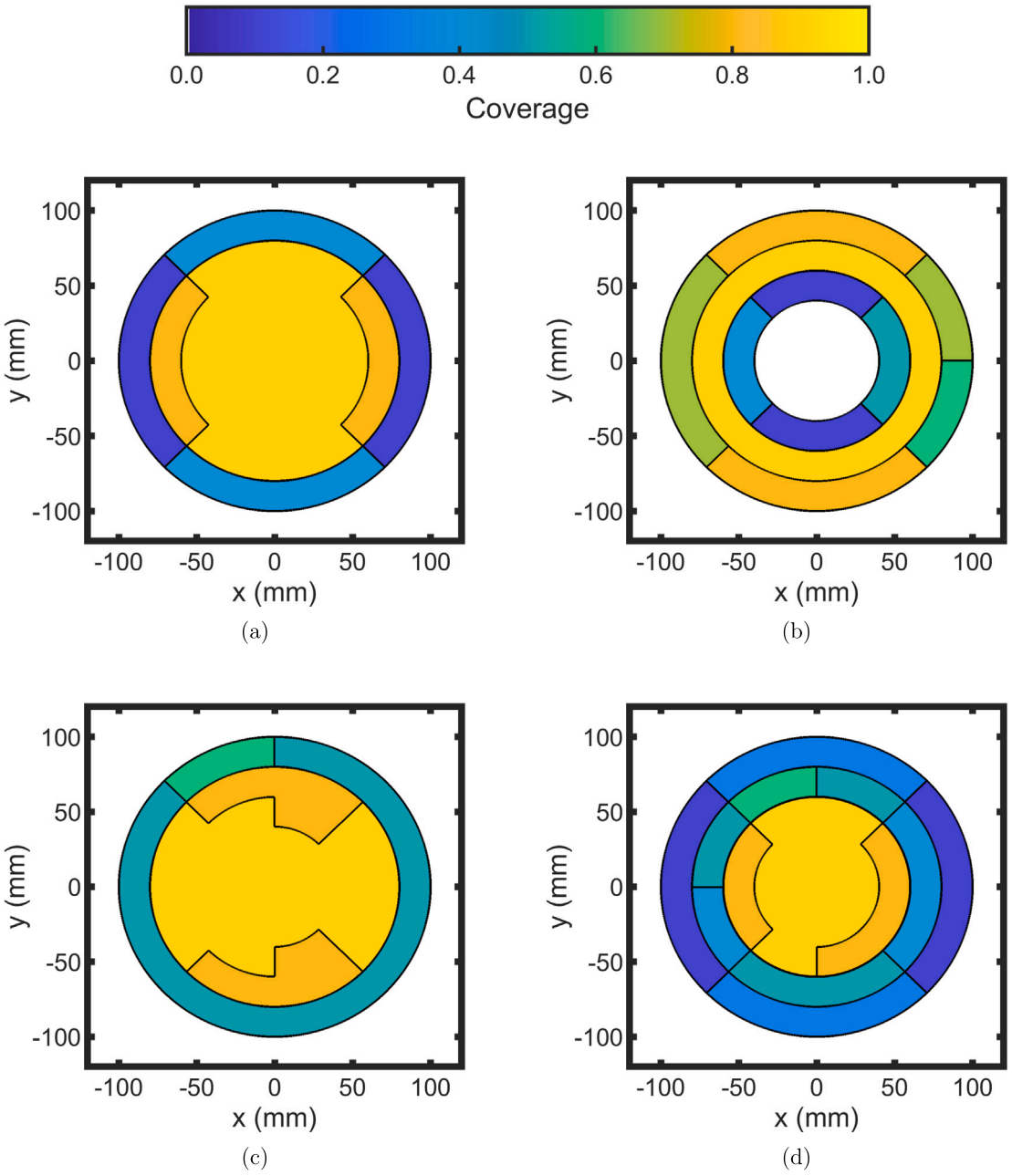

Fig. 8. Comparison of contour plots of various reactor configurations, (a) Single, (b) Ring, (c) Multi, and (d) Combined, to study the spatial behavior of the surface coverage of the terminated Step A product for a 40-partitioned substrate at a time of 0.3 s.

that contour plots are a good metric for identifying efficient reactor designs. Specifically, they are able to compare and determine what reactor design has a better transient pressure profile. Additionally, Figs. 8–10 illustrate slight asymmetric coverage characteristics that are attributed to numerical error from the first-order implicit method utilized to solve the transient transport equations and randomness that originates from the random number used for reaction selection and time advancement in the kMC simulation. While previous work (Tom et al., 2023a) has demonstrated that the discrete feed reactor models produce uniform flow characteristics, the flow is not a significant indicator of coverage uniformity due to the potential of steric hindrance effects and molecular collisions that impede reagent adsorption.

It is also important to point out that the contour plots provide an important insight; there is a significant difference between the developed pressure profiles at 3.0 s and the initial pressure profiles for each reactor design. The former pressure profiles were examined in a previous work, which concluded that the combined-shaped reactor was the best reactor due to its optimal pressure fields (Tom et al., 2023a). However, the results presented in Table 4 differ substantially, as they show that the multi-shaped reactor is the best reactor with the shortest minimum process time. One explanation that accounts for both of these observances is that the pressure fields at the initial stages of the process play a pivotal role in the overall efficacy of the reactor design. The coverage evolution at the fringes of the substrate, which plays the greatest role in determining the minimum process time, is a function of the entire pressure profile evolution. Thus, it is important for the ideal reactor design to also quickly reach these fringe areas. While the combined-shaped reactor may have more even pressure profiles, the multi-shaped reactor evidently distributes the reagent quicker.

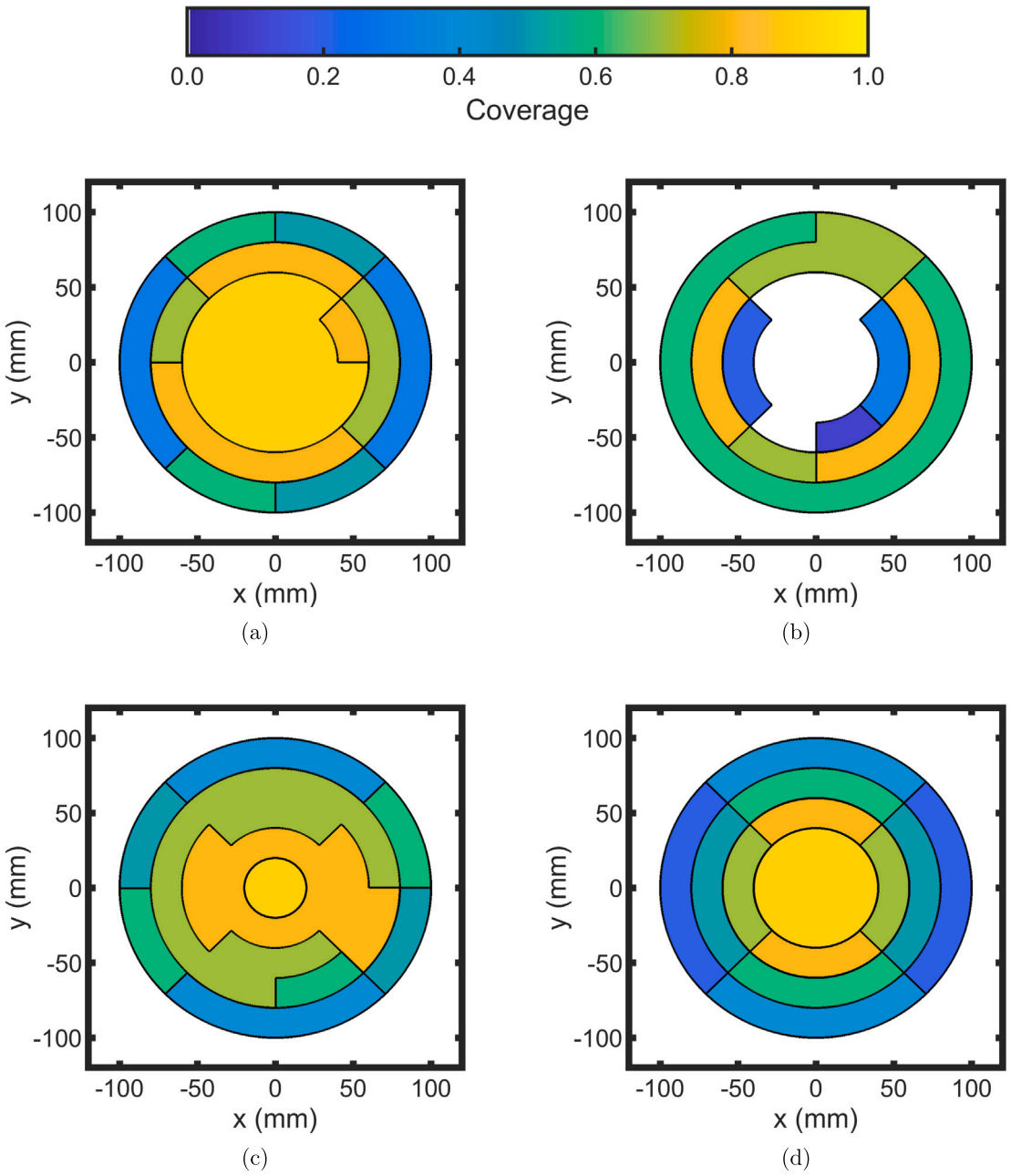

Fig. 9. Comparison of contour plots of various reactor configurations, (a) Single, (b) Ring, (c) Multi, and (d) Combined, to study the spatial behavior of the surface coverage of the terminated Step B product for a 40-partitioned substrate at a time of 0.8 s.

The minimum process time for each reactor design is defined as the time required to reach 99.9% coverage for each individual reaction. Overall, the multi-shaped reactor has the best results with a total minimum process time of 3.638 s. To understand why this reactor design performs the best, the contour plot is instrumental in revealing the fact that this design spread the precursor to the remote parts of the wafer the quickest out of the examined designs. This detail explains why, even though the combined-shaped reactor has the most even pressure profiles, it ranks third in terms of minimum process time.

6. Conclusions

In this work, a multiscale computational fluid dynamics (CFD) model of an area-selective atomic layer deposition (AS-ALD) reactor was developed to study the spatiotemporal progression of surface coverage to allow the scale-up of AS-ALD processes. To improve the accuracy of the multiscale CFD model, this work expanded on the mesoscopic, kinetic Monte Carlo (kMC) algorithm by considering the role of the number of unoccupied atomic sites on the time progression computation. Findings suggested that surface coverage temporal progressions were similar to results conducted in prior work and were able to resolve the ambiguity of surface coverage from prior kMC methodology. Additionally, various reactor injection geometries for the macroscopic CFD simulation were constructed to enhance the mass transfer of reagent on the surface of a semiconductor substrate and study their impact on the process time required to obtain full surface coverage and the uniformity of the surface coverage with respect to time. Results indicated that two reactor models, the multi and the single, required minimal processing time and were characterized by homogeneous, temporal surface coverage.

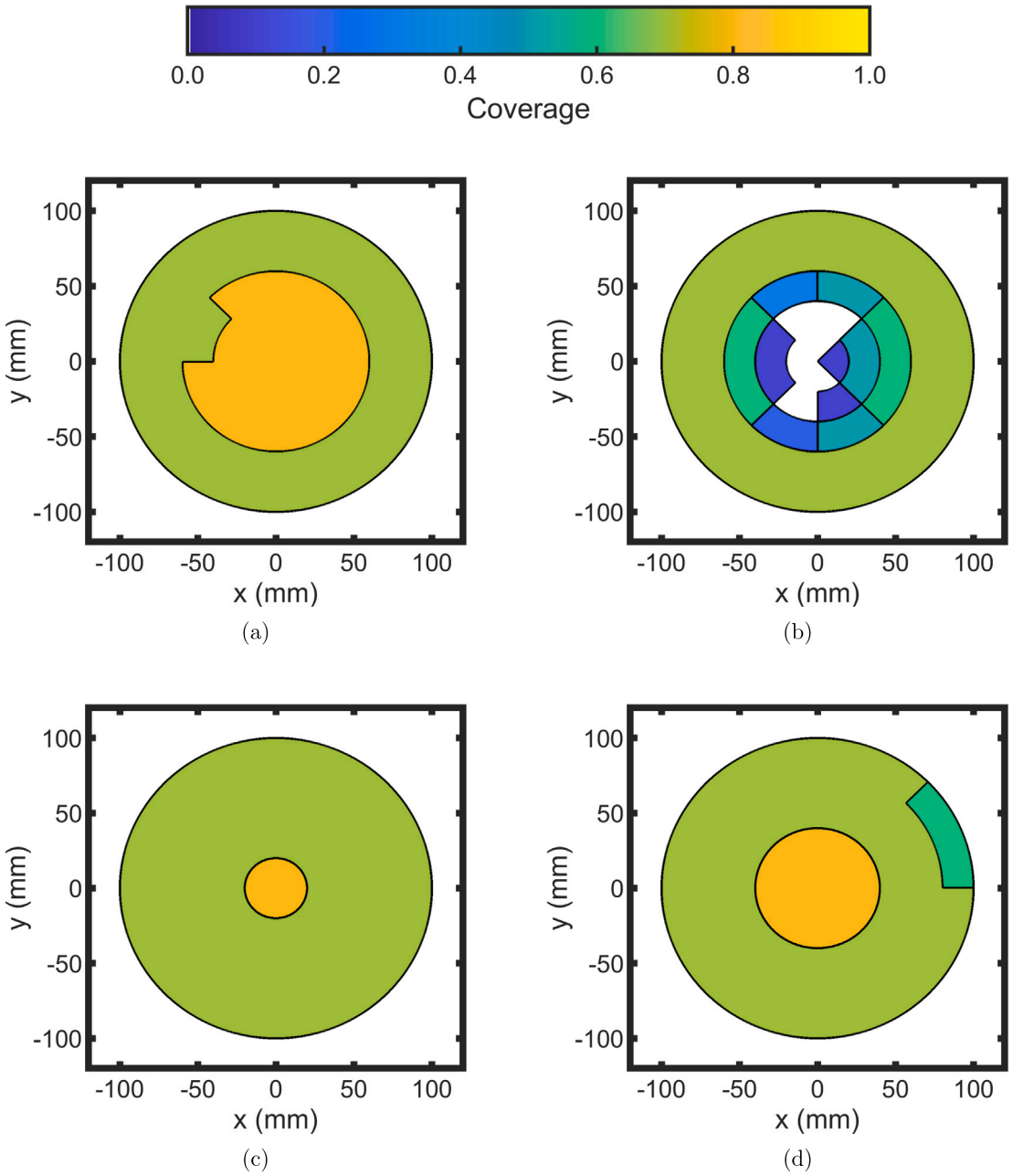

Fig. 10. Comparison of contour plots of various reactor configurations, (a) Single, (b) Ring, (c) Multi, and (d) Combined, to study the spatial behavior of the surface coverage of the terminated Step C product for a 40-partitioned substrate at a time of 1.0 s.

Declaration of competing interest

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

Acknowledgments

Financial support from the National Science Foundation is gratefully acknowledged. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group.

References

ANSYS, 2022a. Ansys Fluent Customization Manual. ANSYS Inc., Canonsburg, PA.
ANSYS, 2022b. Ansys Fluent Theory Guide. ANSYS Inc., Canonsburg, PA.
ANSYS, 2022c. Ansys Fluent User's Guide. ANSYS Inc., Canonsburg, PA.
Bortz, A.B., Kalos, M.H., Lebowitz, J.L., 1975. A new algorithm for Monte Carlo simulation of ising spin systems. J. Comput. Phys. 17, 10–18.
Cheimarios, N., To, D., Kokkoris, G., Memos, G., Boudouvis, A.G., 2021. Monte Carlo and kinetic Monte Carlo models for deposition processes: A review of recent works. Front. Phys. 9, 631918.
Chen, T.-C.T., Chiu, M.-C., 2022. Evaluating the sustainability of smart technology applications in healthcare after the COVID-19 pandemic: A hybridising subjective and objective fuzzy group decision-making approach with explainable artificial intelligence. Digit. Health 8, 20552076221136381.
Chien, C.-F., Kuo, H.-A., Lin, Y.-S., 2022. Smart semiconductor manufacturing for pricing, demand planning, capacity portfolio and cost for sustainable supply chain management. International Journal of Logistics Research and Applications Int. J. Logist. Res. Appl., in press.
Christofides, P.D., Armaou, A., 2006. Control and optimization of multiscale process systems. Comput. Chem. Eng. 30, 1670–1686.
Cremers, V., Puurunen, R.L., Dendooven, J., 2019. Conformality in atomic layer deposition: Current status overview of analysis and modelling. Appl. Phys. Rev. 6, 021302.

DeVita, J.P., Sander, L.M., Smereka, P., 2005. Multiscale kinetic Monte Carlo algorithm for simulating epitaxial growth. Phys. Rev. B 72, 205421.

Dobkin, D.M., Zuraw, M.K. (Eds.), 2003. Principles of Chemical Vapor Deposition, Vol. 1. Springer Dordrecht.

Elers, K.-E., Blomberg, T., Peussa, M., Aitchison, B., Haukka, S., Marcus, S., 2006. Film uniformity in atomic layer deposition. Chem. Vapor Depos. 12, 13–24.

Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22, 403–434.

Giménez, M.C., Del Pópolo, M.G., Leiva, E.P.M., 2002. Kinetic Monte Carlo study of electrochemical growth in a heteroepitaxial system. Langmuir 18, 9087–9094.

Jansen, A. (Ed.), 2012. An Introduction to Kinetic Monte Carlo Simulations of Surface Reactions, Vol. 1. Academic Press, pp. 38–119.

Kim, J., Chakrabarti, K., Lee, J., Oh, K.-Y., Lee, C., 2003. Effects of ozone as an oxygen source on the properties of the Al_2O_3 thin films prepared by atomic layer deposition. Mater. Chem. Phys. 78, 733–738.

Kim, D.S., Kim, J.B., Ahn, D.W., Choe, J.H., Kim, J.S., Jung, E.S., Pyo, S.G., 2023. Atomic layer etching applications in nano-semiconductor device fabrication. Electron. Magn. Photon. 19, 424–441.

Kim, H.G., Kim, M., Gu, B., Khan, M.R., Ko, B.G., Yasmeen, S., Kim, C.S., Kwon, S.-H., Kim, J., Kwon, J., Jin, K., Cho, B., Chun, J.S., Shong, B., Lee, H.-B.-R., 2020. Effects of Al precursors on deposition selectivity of atomic layer deposition of Al_2O_3 using ethanethiol inhibitor. Chem. Mater. 32, 8921–8929.

Klement, P., Anders, D., Gümbel, L., Bastianello, M., Michel, F., Schörmann, J., Elm, M.T., Heiliger, C., Chatterjee, S., 2021. Surface diffusion control enables tailored-aspect-ratio nanostructures in area-selective atomic layer deposition. ACS Appl. Mater. Interfaces 13, 19398–19405.

Leslie, M., 2022. Pandemic scrambles the semiconductor supply chain. Engineering 9, 10–12.

Lin, S.-C., Wang, C.-C., Tien, C.-L., Tung, F.-C., Wang, H.-F., Lai, S.-H., 2023. Fabrication of aluminum oxide thin-film devices based on atomic layer deposition and pulsed discrete feed method. Micromachines 14.

Mameli, A., Merkx, M.J.M., Karasulu, B., Roozeboom, F., Kessels, W.E.M.M., Mackus, A.J.M., 2017. Area-selective atomic layer deposition of SiO_2 using acetylacetone as a chemoselective inhibitor in an ABC-type cycle. ACS Nano 11, 9303–9311.

Maroudas, D., 2000. Multiscale modeling of hard materials: Challenges and opportunities for chemical engineering. AIChE J. 46, 878–882.

Merkx, M.J.M., Sandoval, T.E., Hausmann, D.M., Kessels, W.M.M., Mackus, A.J.M., 2020. Mechanism of precursor blocking by acetylacetone inhibitor molecules during area-selective atomic layer deposition of SiO_2. Chem. Mater. 32, 3335–3345.

Moore, G.E., 1998. Cramming more components onto integrated circuits. Proc. IEEE 86, 82–85.

Oh, I.-K., Sandoval, T.E., Liu, T.-L., Richey, N.E., Bent, S.F., 2021. Role of precursor choice on area-selective atomic layer deposition. Chem. Mater. 33, 3926–3935.

Radamson, H., 2018. 1 - Basics of metal–oxide–semiconductor field-effect transistor (MOSFET). In: Radamson, H.H., Luo, J., Simoen, E., Zhao, C. (Eds.), CMOS Past, Present and Future. Woodhead Publishing, pp. 1–17.

Richard, C., 2023. Understanding Semiconductors: A Technical Guide for Non-Technical People. A Press, Berkeley, CA.

Tom, M., Wang, H., Ou, F., Yun, S., Orkoulas, G., Christofides, P.D., 2023a. Computational fluid dynamics modeling of a discrete feed atomic layer deposition reactor: Application to reactor design and operation. Comput. Chem. Eng. 178, 108400.

Tom, M., Yun, S., Wang, H., Ou, F., Orkoulas, G., Christofides, P.D., 2023b. Multiscale modeling of spatial area-selective thermal atomic layer deposition. In: Kokossis, A.C., Georgiadis, M.C., Pistikopoulos, E. (Eds.), Proceedings of 33rd European Symposium on Computer Aided Process Engineering. In: Computer Aided Chemical Engineering, vol. 52, Athens, Greece, pp. 71–76.

Wehinger, G.D., Ambrosetti, M., Cheula, R., Ding, Z.-B., Isoz, M., Kreitz, B., Kuhlmann, K., Kutscherauer, M., Niyogi, K., Poissonnier, J., Réocreux, R., Rudolf, D., Wagner, J., Zimmermann, R., Bracconi, M., Freund, H., Krewer, U., Maestri, M., 2022. Quo vadis multiscale modeling in reaction engineering? – A perspective. Chem. Eng. Res. Des. 184, 39–58.

Yuan, C.Y., Sheng, Y., 2010. Sustainable scale-up studies of atomic layer deposition for microelectronics manufacturing. In: Proceedings of the 2010 IEEE International Symposium on Sustainable Systems and Technology. Arlington, VA, USA, pp. 1–6.

Yun, S., Ou, F., Wang, H., Tom, M., Orkoulas, G., Christofides, P.D., 2022a. Atomistic-mesoscopic modeling of area-selective thermal atomic layer deposition. Chem. Eng. Res. Des. 188, 271–286.

Yun, S., Tom, M., Ou, F., Orkoulas, G., Christofides, P.D., 2022b. Multiscale computational fluid dynamics modeling of thermal atomic layer etching: Application to chamber configuration design. Comput. Chem. Eng. 161, 107757.

Yun, S., Wang, H., Tom, M., Ou, F., Orkoulas, G., Christofides, P.D., 2023. Multiscale CFD modeling of area-selective atomic layer deposition: Application to reactor design and operating condition calculation. Coatings 13, 558.

Zhang, J., Li, Y., Cao, K., Chen, R., 2022. Advances in atomic layer deposition. Nanomanufact. Metrol. 5, 191–208.

Zhang, L., Wang, S., Ma, C., He, J., Xu, C., Ma, Y., Ye, Y., Liang, H., Chen, Q., Chan, M., 2012. Gate underlap design for short channel effects control in cylindrical gate-all-around MOSFETs based on an analytical model. IETE Tech. Rev. 29, 29–35.