TC-FNO: Fourier Neural Operator with Learnable Time-Step Embedding for 3D Lithography Exposure Simulation

Yuqian Zang[1], Yiqing Wang[1], Silin Chen[1], Jiale Li[1], Zhili Xu[1], Yufeng Zhang[1],
Hao Zhang[3†], Chia-Yen Li[4†], Ningmu Zou[1,2†]
[1]School of Integrated Circuits, Nanjing University, Suzhou, China
[2]Interdisciplinary Research Center for Future Intelligent Chips (Chip-X), Nanjing University, Suzhou, China
[3]Suzhou Laboratory, Suzhou, China
[4]NexChip Co., Ltd., China

***Abstract*—In chemically amplified photoresist, optical exposure transforms the spatial light intensity distribution into a three-dimensional photoacid distribution, which in turn drives the post-exposure bake (PEB) and development processes. Rigorous exposure simulation is typically formulated as solving a system of coupled partial differential equations on a dense three-dimensional mesh. However, solving these equations is computationally expensive and can become a bottleneck in photoresist simulation and optimization. Inspired by the physical structure of the exposure process, we propose a time-conditioned Fourier neural operator as a neural PDE surrogate for exposure simulation. The key innovation lies in explicitly encoding the time-step size Δt to constrain the learned evolution operator, since discrete exposure updates inherently depend on Δt. This model learns the mapping from the current three-dimensional photoacid field and light intensity field to the photoacid field at the next time step. Compared to traditional numerical solvers, it can achieve accurate prediction of three-dimensional photoacid evolution with significantly reduced runtime. This method provides efficient exposure modeling support for photoresist simulation, illumination/mask optimization, and layout co-optimization.**

***Index Terms*—Lithography simulation, photoacid generation, exposure modeling, Fourier Neural Operator, neural PDE solver, 3D field prediction.**

I. Introduction

OPTICAL lithography remains the workhorse for advanced semiconductor manufacturing, enabling high-volume production of integrated circuits at ever-shrinking feature sizes [1]. As the critical dimension (CD) continues to approach the resolution limit of projection systems, computational lithography has become indispensable for predicting pattern fidelity and for guiding layout optimization, source–mask co-optimization, and advanced process control [2]–[4]. A key requirement for these applications is an accurate and efficient modeling of the full photoresist process, which converts a given mask layout and illumination setting into the final developed resist profile [1].

†Corresponding authors: nzou@nju.edu.cn, zhangh@szlab.ac.cn, chi-ayenli@nexchip.com.cn. This work was supported in part by the National Natural Science Foundation of China (62341408).

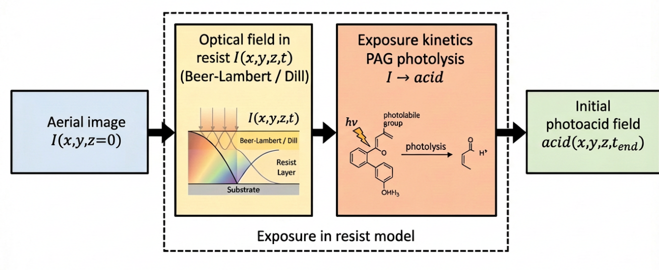


Fig. 1. Schematic of the exposure-in-resist model considered in this work. The aerial image $I(x, y, z = 0)$ from the optical imaging module is first propagated into the resist to obtain the optical field $I(x, y, z, t)$ (e.g., via Beer–Lambert or Dill formulations [1], [5]), which then drives PAG photolysis kinetics $I \rightarrow$ acid and yields the initial 3D photoacid field $\mathrm{acid}(x, y, z, t_{\mathrm{end}})$ at the end of exposure.

A typical resist simulation flow consists of three tightly coupled stages: exposure, and post-exposure bake (PEB) and development [1], [6]. Optical imaging computes the aerial image in the resist from the mask and the projection system. Exposure converts this aerial image into the initial three-dimensional (3D) distribution of chemically amplified species (e.g., photoacid), while the post-exposure bake (PEB) and development modules model subsequent reaction–diffusion dynamics and material dissolution to obtain the latent and developed images [6], [8]–[10]. The exposure stage is critical because it provides the initial condition for PEB and development, so exposure errors can propagate downstream. In addition, exposure simulation is time-dependent and typically implemented by discrete time stepping, where the choice of Δt affects the discrete evolution operator and the resulting 3D photoacid distribution. The corresponding exposure-in-resist pipeline is summarized in Fig. 1.

Rigorous exposure simulation is typically formulated as coupled PDEs describing the spatio-temporal evolution of photoacid and related species under optical illumination [1], [5]. These PDEs are solved on dense 3D grids with appropriate

initial/boundary conditions, often via finite-difference or finite-volume schemes. Although accurate, such solvers are computationally expensive, especially for full-chip layouts, large process windows, or repeated evaluation in optimization loops [1], [4]. This cost is largely dominated by repeatedly solving tightly coupled 3D multi-physics models on fine spatial grids over many time steps and parameter settings. In addition, wave-optical effects such as interference and standing-wave patterns induced by reflections at material interfaces require fine resolution along the resist thickness direction, further increasing the computational cost. As a result, many industrial flows still use compact or empirical exposure models that trade accuracy for speed, limiting resist simulation fidelity [1].

Recent years have seen growing interest in applying machine learning (ML) to computational lithography, where CNN-based models have been explored for fast resist modeling, OPC/ILT, source–mask optimization, and hotspot detection [11]–[13]. Most of these methods formulate lithography as supervised image-to-image regression from mask/aerial images to resist metrics, while rarely modeling the inherently 3D and time-dependent exposure dynamics. As a result, many approaches rely on 2D representations or collapse exposure and post-exposure effects into a black-box surrogate, which can limit interpretability and generalization across process conditions.

From a physical viewpoint, exposure can be cast as learning an evolution operator that maps the current 3D photoacid and intensity fields to the next time step. Neural operators, especially Fourier Neural Operators (FNOs), provide an effective way to approximate PDE solution operators directly in function space [16], [18], offering a flexible surrogate that can generalize across resolutions. A critical detail is that this evolution operator is not unique: it is parameterized by the time increment Δt used in the discrete exposure update. If Δt is not explicitly provided, a vanilla FNO is forced to approximate a single implicit update rule, which leads to operator mismatch when temporal transitions vary and consequently degrades accuracy, especially near CD-critical edges.

In this paper, we propose a time-conditioned Fourier Neural Operator with explicit Δt conditioning, called TC-FNO, to learn a family of discrete exposure-evolution operators. TC-FNO maps the current 3D photoacid and intensity fields to the next-step photoacid field using a time embedding and is trained on samples generated by a commercial exposure simulator.

The main contributions of this work are summarized as follows:

- We formulate 3D photoacid evolution during exposure as learning a time-stepping operator, and construct a time-series field dataset that pairs (acid_t, I_t) with $\mathrm{acid}_{t+\Delta t}$ on dense 3D grids, Δt is retained as an explicit conditioning variable to select the corresponding operator.
- We propose a time-conditioned Fourier Neural Operator that jointly models spatial correlations via spectral convolutions and temporal dynamics through time-conditioned embeddings, providing a physically motivated neural surrogate for the underlying exposure PDEs.

II. Preliminaries

This section summarizes the PDE formulation of the exposure process, introduces the operator-learning viewpoint and Fourier Neural Operators, and specifies the error metrics used in this work.

A. Exposure PDE and Time Stepping

We consider a chemically amplified resist on a 3D domain $\Omega \subset \mathbb{R}^3$ with $\mathbf{x} = (x, y, z) \in \Omega$ and $t \in [0, T]$. An i-line chemically amplified system [7] typically consists of a polymer matrix, a bound photoacid (PAC), and a solvent [1], [6]. The normalized PAC concentration $m(x, z, t)$ (latent image) satisfies $m(x, z, 0) = 1$. The intensity field $I(x, z, t)$ is obtained from the top-surface aerial image $I(x, y, z = 0)$ via Beer–Lambert/Dill-type attenuation models [1], [5].

A classical description of exposure in this setting is the Dill model [5]. The PAC is depleted according to a first-order kinetic law, while the intensity decays along depth as

$$\frac{\partial m(x,y,z,t)}{\partial t} = -\,C\,I(x,y,z,t)\,m(x,y,z,t), \tag{1}$$

$$\frac{\partial I(x,y,z,t)}{\partial z} = -\big[A\,m(x,y,z,t) + B\big]\,I(x,y,z,t), \tag{2}$$

where A, B, and C are the Dill coefficients of the resist. Without additional assumptions, this coupled system cannot be solved in closed form, and a common strategy is to decompose the total exposure time t_{expo} into small time steps Δt [1]. Over each step, one alternately updates m and I by integrating the kinetic and attenuation equations. For example, the PAC update can be written as

$$m(x,z,t+\Delta t) = m(x,z,t)\,\exp\big(-\,C\,I(x,z,t)\,\Delta t\big), \tag{3}$$

while keeping I quasi-stationary within the small interval. The resulting evolution of the chemically amplified species can then be summarized compactly as a reaction-type PDE for a state variable $a(\mathbf{x}, t)$ that represents the normalized photoacid concentration [1], [6].

$$\frac{\partial a(\mathbf{x},t)}{\partial t} = f\big(a(\mathbf{x},t), I(\mathbf{x},t); \lambda\big), \quad \mathbf{x} \in \Omega,\ t \in [0,T], \tag{4}$$

where $f(\cdot)$ encodes the effective photochemical mechanism and λ collects the material parameters (e.g. the Dill coefficients).

For numerical simulation, (4) is discretized on a Cartesian grid $\{\mathbf{x}_{i,j,k}\}$ and time instances $t_0 < t_1 < \cdots < t_{N_t}$ with step size $\Delta t = t_{n+1} - t_n$. Denoting $a^n_{i,j,k} \approx a(\mathbf{x}_{i,j,k}, t_n)$ and $I_{i,j,k} \approx I(\mathbf{x}_{i,j,k}, t_n)$, an explicit time integrator can be written as

$$a^{n+1}_{i,j,k} = a^n_{i,j,k} + \Delta t\, f\big(a^n_{i,j,k}, I_{i,j,k}; \lambda\big), \tag{5}$$

which we view as repeatedly applying a deterministic time–stepping operator $\mathcal{T}_{\Delta t} : (a^n, I) \mapsto a^{n+1}$.

It should be emphasized that the discrete operator $\mathcal{T}_{\Delta t}$ is not invariant to the time-step size. Different choices of Δt yield different numerical realizations of the same continuous-time exposure dynamics, even when the governing PDE is

unchanged. Therefore, a surrogate that does not explicitly encode Δt is effectively constrained to approximate a single frozen operator, which can degrade accuracy and robustness when the temporal discretization varies.

B. Neural operator-based PDE solver

Recent years have seen rapid progress in neural-network-based PDE solvers, which can be broadly grouped into three categories. Physics-informed surrogates embed governing equations and boundary/initial conditions in the loss [14], they are mesh-free and yield continuous solutions, but typically require a new optimization per PDE instance, which is costly for large-scale exposure simulation. Finite-dimensional models learn mappings between discretized input/output fields on a fixed grid, with sufficient data, they can approximate solutions, but their resolution dependence limits generalization across meshes and domains.

Neural operators address these limitations by learning a mesh-free solution operator for a family of PDEs with a single set of parameters [15], [16]. Instead of finite-length vectors, they act on functions over a bounded domain and learn the operator in function space. Among them, Fourier Neural Operators (FNOs) are effective and efficient due to kernel integration implemented in the Fourier domain [18].

Concretely, given a feature field $v_k : \Omega \to \mathbb{R}^C$ at layer k, a 3D FNO layer updates it according to

$$v_{k+1}(\mathbf{x}) = \sigma\Big(W v_k(\mathbf{x}) + \mathcal{F}^{-1}\big(R_k(\boldsymbol{\xi}) \cdot \mathcal{F}[v_k](\boldsymbol{\xi})\big)\Big), \quad (6)$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the 3D Fourier and inverse Fourier transforms, $\boldsymbol{\xi}$ is the frequency variable, $R_k(\boldsymbol{\xi})$ is a learnable complex-valued tensor defined on a truncated set of low-frequency modes, W is a local linear transformation (applied pointwise in the spatial domain), and $\sigma(\cdot)$ is a nonlinear activation function. By stacking several layers of the form (6), one obtains a deep neural operator with an effectively global receptive field and a compact spectral parameterization.

In this work, we adopt a time-conditioned 3D FNO as the neural operator-based PDE solver for the exposure stage, learning the evolution operator that advances the 3D photoacid field in chemically amplified resists.

C. Problem Formulation

To assess how well the learned surrogate reproduces the commercial exposure simulator, we evaluate both volumetric discrepancies between 3D fields and pattern-level errors on developed profiles.

Let G_{pred} and G_{label} denote the predicted and ground-truth 3D distributions of a scalar quantity of interest on a grid of size $N_x \times N_y \times N_z$. For any tensor D defined on this grid, we use the Frobenius norm

$$\|D\|_F = \Big(\sum_{i,j,k} D_{i,j,k}^2\Big)^{1/2}. \quad (7)$$

Root Mean Squared Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\big\|G_{\text{pred}} - G_{\text{label}}\big\|_F^2}, \qquad n = N_x N_y N_z. \quad (8)$$

Normalized RMSE (NRMSE): To obtain scale-invariant measures, we additionally use normalized RMSE. The basic form is

$$\text{NRMSE} = \frac{\big\|G_{\text{pred}} - G_{\text{label}}\big\|_F}{\big\|G_{\text{label}}\big\|_F}, \quad (9)$$

where $\|\cdot\|_F$ denotes the Frobenius norm, given by $\|D\|_F = \Big(\sum_i \sum_j D_{ij}^2\Big)^{1/2}$.

Critical Dimension (CD) Error: In addition to field-based metrics, we also evaluate the error in lithography-relevant critical dimensions extracted from the simulated profiles. For each test pattern and measurement direction $\alpha \in \{x, y\}$, let $CD_{\alpha,\text{pred}}^{(i)}$ and $CD_{\alpha,\text{label}}^{(i)}$ denote the predicted and reference CD values at the i-th sampling location, $i = 1, \ldots, n_\alpha$. The CD error is defined as

$$\text{CD}_\alpha\text{-Error} = \sqrt{\frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha}(CD_{\alpha,\text{pred}}^{(i)} - CD_{\alpha,\text{label}}^{(i)})^2}, \quad \alpha \in \{x, y\} \quad (10)$$

Based on the above PDE formulation, neural-operator framework, and error metrics, we define the exposure modeling problem studied in this work as follows:

Given commercial exposure simulations that provide the 3D photoacid distribution and corresponding developed profiles under various mask patterns and process conditions, the objective is to design and train a neural operator to predict the 3D photoacid distribution of unseen test patterns with low RMSE, NRMSE, and CD errors relative to the commercial simulator, while achieving sufficiently low runtime for large-scale lithography simulation.

III. Method

In this section, we present the proposed TC-FNO framework for exposure modeling in chemically amplified resists. We first describe how high-fidelity simulation results are organized into a supervised learning dataset, then introduce the TC-FNO architecture used to approximate the exposure evolution, and finally detail the training objective and optimization strategy.

A. Data Representation and Supervision

High-fidelity reference data are generated by a commercial exposure simulator for a set of representative mask patterns and a sequence of exposure times $t_0, \ldots, t_{N_t}$. For each cell c and each pair of consecutive time instants (t_n, t_{n+1}) within the exposure window, we record

- the photoacid field $a_c(\mathbf{x}, t_n)$,
- the corresponding optical intensity field $I_c(\mathbf{x})$,
- the next-step photoacid field $a_c(\mathbf{x}, t_{n+1})$.

All continuous fields are sampled on a regular 3D Cartesian grid and, if necessary, downsampled to a common resolution $X \times Y \times Z$. Each channel is standardized using global statistics computed over the training set so that the dynamic range is comparable across cells and time steps. For a given time pair (t_n, t_{n+1}), the network input is a two-channel tensor

$$\mathbf{x} = \big[a(\cdot, t_n),\, I(\cdot)\big] \in \mathbb{R}^{2\times X\times Y\times Z}, \quad (11)$$

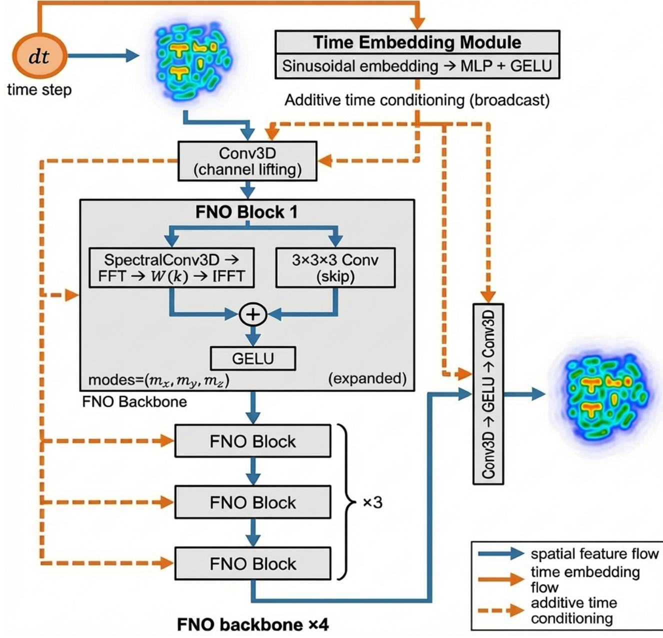


Fig. 2. Time-conditioned 3D FNO architecture for photoacid field forecasting.

and the supervision target is the subsequent photoacid field

$$\mathbf{y} = a(\cdot, t_{n+1}) \in \mathbb{R}^{1\times X\times Y\times Z}. \tag{12}$$

The time increment $\Delta t = t_{n+1} - t_n$ is retained as an additional scalar conditioning variable.

Aggregating samples over all cells and admissible time pairs yields a large spatio–temporal dataset of the form $(\mathbf{x}, \Delta t) \mapsto \mathbf{y}$, which directly discretizes the exposure evolution operator $\mathcal{T}_{\Delta t} : (a^n, I) \mapsto a^{n+1}$.

To avoid pattern leakage, we partition the dataset into disjoint training/validation/test splits at the *mask* level, so that layouts in the test set are never seen during training. Within the training cells, we sample time pairs for optimization and use a held-out subset of time pairs for validation. All results reported in Section IV are evaluated on the held-out test cells across the full exposure window.

B. Time-Conditioned 3D FNO Architecture

We instantiate $\mathcal{G}_\theta$ as a 3D Fourier Neural Operator acting on the (z, y, x) grid. Given a pair $(\mathbf{x}, \Delta t)$, the model outputs a prediction $\hat{\mathbf{y}}$ that approximates the next-time-step photoacid field. An overview of the architecture and data flow is shown in Fig. 2.

- **Input lifting.** The input tensor $\mathbf{x}$ is first embedded by a 3D convolutional layer (kernel size $3 \times 3 \times 3$, C output channels), which maps the two physical channels into a C-dimensional feature representation. By linearly mixing the photoacid and intensity fields within local spatial neighborhoods, this layer embeds the raw inputs into a shared latent representation that serves as the domain of the subsequent operator layers.
- **Spectral backbone.** The lifted feature map is processed by a stack of L 3D FNO blocks, each implementing the spectral update rule in (6) with a limited number of retained Fourier modes along each spatial dimension. The spectral convolution captures long-range spatial correlations, while an additional $3\times3\times3$ convolutional skip path and residual connections enhance local modeling capacity and improve training stability.
- **Time embedding and conditioning.** The scalar time increment Δt is mapped to a time embedding $e_t = \phi(\Delta t) \in \mathbb{R}^C$ via a sinusoidal encoding followed by a small multilayer perceptron with GELU activation. This embedding is broadcast over the spatial grid and injected into each FNO block through an additive residual pathway, thereby conditioning the learned operator on the time step while preserving the spatial structure of the fields.

TABLE I
COMPARISON OF DIFFERENT ALGORITHMS (ACID ACCURACY, CD ERROR, AND RUNTIME).

Algorithm	Acid		CD Error / nm		Runtime/s
	RMSE/E-3	NRMSE/%	X	Y	
DeepCNN [19]	7.24	12.40	4.250	4.211	3.75
CGAN [17]	9.16	13.92	4.642	5.168	2.54
FNO [18]	6.56	11.23	3.785	4.343	3.98
TC-FNO	4.46	7.63	1.119	1.028	0.93

From an operator-learning perspective, the time embedding does not merely serve as auxiliary input, but enables the network to represent a continuum of exposure evolution operators parameterized by Δt. Without such conditioning, the Fourier backbone is forced to average over heterogeneous temporal transitions in the training data, which degrades its ability to resolve stiff photoacid dynamics and fine-scale temporal variations.

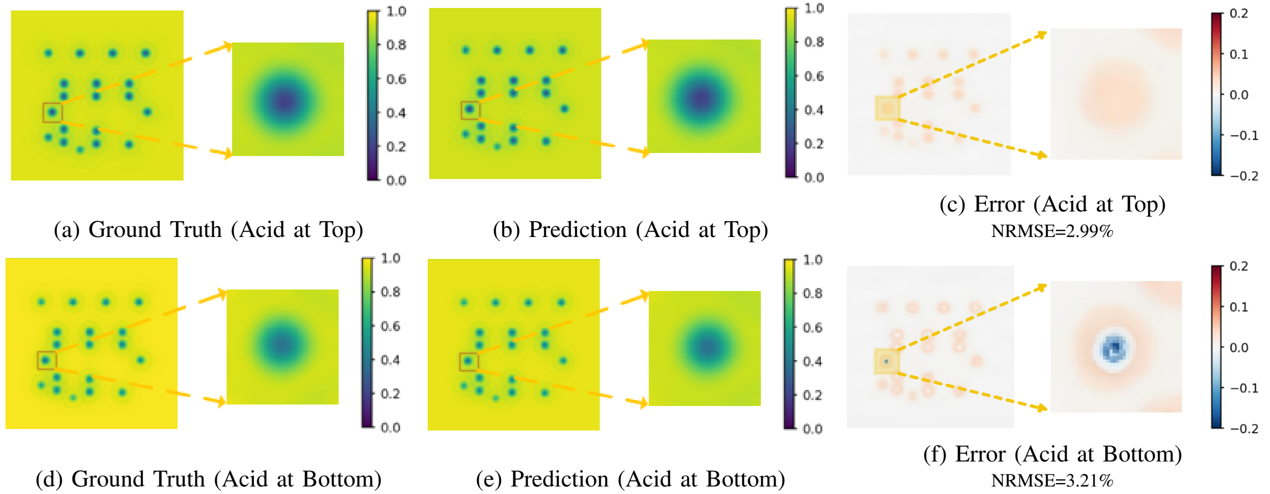
C. Training Objective and Optimization

The time-conditioned 3D FNO is trained in a supervised manner to regress from $(\mathbf{x}, \Delta t)$ to the reference 3D photoacid field $\mathbf{y}$. For a mini-batch of size B with $n = XYZ$ voxels per sample, let $\hat{\mathbf{y}}^{(b)}$ and $\mathbf{y}^{(b)}$ denote the predicted and reference fields, and $r_i^{(b)} = \hat{y}_i^{(b)} - y_i^{(b)}$ the residual at voxel i of sample b. We adopt a mixed L^2–L^∞ objective:

$$\mathcal{L}_{\text{train}}(\theta) = \frac{1}{Bn}\sum_{b=1}^{B}\sum_{i=1}^{n}\left(r_i^{(b)}\right)^2 + \lambda_{\max} \max_{1\le b\le B}\ \max_{1\le i\le n}\left(r_i^{(b)}\right)^2, \tag{13}$$

where $\lambda_{\max} > 0$ balances the average voxel-wise error and the worst-case local error.

The parameters θ are optimized using the Adam optimizer [20] with a cosine-annealed learning rate schedule. Network width and depth, the number of retained Fourier modes, batch size, and training budget are kept fixed across all experiments. In addition to the training loss, the RMSE,



(a) Ground Truth (Acid at Top) (b) Prediction (Acid at Top) (c) Error (Acid at Top) NRMSE=2.99%

(d) Ground Truth (Acid at Bottom) (e) Prediction (Acid at Bottom) (f) Error (Acid at Bottom) NRMSE=3.21%

Fig. 3. Example of the predicted acid latent image for two vertical distribution types (*Acid at Top* and *Acid at Bottom*). For each row, the three panels show the ground-truth acid field, the prediction of the proposed time-conditioned FNO, and the error map ($G_{\text{pred}} - G_{\text{label}}$). Within each panel, the global view, a zoomed-in center region, and the associated color bar are displayed from left to right.

NRMSE, and CD errors defined in Section II are monitored on the validation set and reported in the experimental results.

IV. Experiments

A. Evaluation Protocol

We evaluate the proposed time-conditioned 3D FNO on the held-out test cells in Section III. The dataset is generated by a commercial exposure simulator on representative layout patterns under multiple process conditions. For each cell and consecutive exposure times, the 3D photoacid field and optical intensity are sampled on a regular grid to form supervised pairs (Section III-B).

All models share the same train/validation split, normalization, and optimization schedule. On each test pattern, the surrogate predicts the next-step 3D photoacid field, we report Acid RMSE/NRMSE (8)–(9), CD errors in x/y from predefined measurement locations (10), and average wall-clock inference runtime over all test instances on the test hardware.

B. Comparison with Baseline Models

Table I summarizes the quantitative comparison between the proposed time-conditioned FNO (TC-FNO) and three representative baselines: a conditional GAN (CGAN) [17], a vanilla 3D FNO (FNO) [18], and a DeepCNN model [19].

In terms of Acid RMSE/NRMSE, TC-FNO achieves the best accuracy. Compared with DeepCNN, RMSE decreases from 7.24×10^{-3} to 4.46×10^{-3} and NRMSE from 12.40% to 7.63% (about 38% reduction). Relative to vanilla FNO, RMSE improves from 6.56×10^{-3} to 4.46×10^{-3} and NRMSE from 11.23% to 7.63% (about 30–35$\%$ reduction). CGAN yields the largest Acid errors, suggesting the adversarial formulation is less effective for fine-scale 3D photoacid evolution in this setting.

CD errors exhibit an even more pronounced gap. In the x-direction, the average CD error of TC-FNO is $1.119\,\text{nm}$, compared with $4.250\,\text{nm}$ for DeepCNN, $3.785\,\text{nm}$ for FNO, and $4.642\,\text{nm}$ for CGAN. In the y-direction, TC-FNO attains an average error of $1.028\,\text{nm}$, whereas the baselines range from $4.211\,\text{nm}$ to $5.168\,\text{nm}$. Overall, the proposed model reduces CD errors by more than a factor of three relative to the best baseline, demonstrating that the improved field-level accuracy translates directly into lithography-relevant pattern fidelity.

Runtime results in Table I show that TC-FNO is also efficient, achieving the shortest inference time ($0.93\,\text{s}$ per prediction), faster than DeepCNN and FNO. CGAN is slightly faster than DeepCNN but still slower than TC-FNO, indicating a favorable accuracy–cost trade-off for large-scale exposure modeling.

In addition to the quantitative metrics reported above, we provide a qualitative comparison in Fig. 3. For two representative acid distribution types (*Acid at Top* and *Acid at Bottom*), each row shows the mask pattern, the ground-truth 3D acid latent image, the prediction of the proposed time-conditioned FNO, and the corresponding error map. Within each panel, the global view and a zoomed-in center region are displayed together with the color bar.

As shown in Fig. 3, the predicted acid fields closely match the ground truth for both distribution types, preserving the overall spot shapes and intensities. Residual errors are small and mainly localized around contact edges. This is consistent with Table I, confirming that the proposed model captures the spatial structure of the acid distribution with low and localized prediction errors.

C. Ablation Study

To understand the contribution of individual design choices in TC-FNO, we conduct an ablation study whose results are reported in Table II.

Effect of time embedding. The pronounced performance gap observed in the ablation study can be directly attributed to the role of Δt in defining the exposure evolution op-

TABLE II
RESULTS OF THE ABLATION STUDY.

Configuration			NRMSE/%	CD Error/nm	
Fourier Branch	Time Embedding	Superior Norm Loss	Acid	X	Y
✓			9.22	3.037	3.644
✓	✓		5.13	1.864	1.794
✓	✓	✓	4.69	1.119	1.028

erator. Configuration 1 uses only the Fourier branch and serves as the baseline. Configuration 2 augments the model with the proposed sinusoidal time embedding and additive conditioning. Introducing time embedding reduces the Acid NRMSE from 9.22% to 5.13%, and the CD errors drop from $(3.037, 3.644)\,\mathrm{nm}$ in (x, y) to $(1.864, 1.794)\,\mathrm{nm}$. These improvements confirm that explicitly encoding the time increment as a conditioning signal is crucial for capturing the temporal evolution of the photoacid field.

Effect of superior norm loss. The variant that further incorporates the superior norm loss emphasizes deviations in high-error regions of the 3D field. This leads to an additional gain: the Acid NRMSE is reduced to 4.69%, and the CD errors are brought down to $(1.119, 1.028)\,\mathrm{nm}$ in (x, y), which matches the best performance reported in Table I. The ablation results demonstrate that both the time embedding mechanism and the enhanced loss formulation are necessary to fully exploit the capacity of the Fourier operator backbone and to attain state-of-the-art accuracy on this task.

V. CONCLUSION

In this work, we proposed a time-conditioned 3D Fourier Neural Operator for exposure modeling in chemically amplified resists. Based on a PDE description of photoacid evolution, we formulate exposure as learning an operator on 3D fields and construct supervised data from rigorous Sentaurus simulations. The model takes the current photoacid and intensity fields together with the time increment to predict the next-step photoacid distribution. As summarized in Table I, TC-FNO improves Acid RMSE/NRMSE over DeepCNN, CGAN, and vanilla 3D FNO, while achieving the lowest CD errors in both x and y and the fastest inference. Table II further verifies the contribution of explicit time embedding and the superior norm loss, particularly for CD accuracy. Future work will study multi-step rollout, coupled exposure–PEB–development modeling, and integration into OPC/ILT pipelines.

REFERENCES

[1] C. A. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. Chichester, West Sussex, England: John Wiley & Sons, Ltd., 2007.

[2] A. Erdmann, T. Fühner, P. Evanschitzky, V. Agudelo, C. Freund, P. Michalak, and D. Xu, "Optical and EUV projection lithography: A computational view," *Microelectronic Engineering*, vol. 132, pp. 21–34, 2015, doi: 10.1016/j.mee.2014.09.011.

[3] D. Flagello and C. A. Mack, "Guest Editorial: Computational lithography," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 8, no. 3, Art. no. 031401, 2009, doi: 10.1117/1.3190214.

[4] D. O. S. Melville, S. D. Slonaker, F. B. Pollard, and D. J. Krafcik *et al.*, "Computational lithography: Exhausting the resolution limits of 193-nm projection lithography systems," *Journal of Vacuum Science & Technology B*, vol. 29, no. 6, Art. no. 06F305, 2011, doi: 10.1116/1.3662090.

[5] F. H. Dill, A. R. Neureuther, J. A. Rothschild, and B. J. Shaw, "Modeling projection printing of positive photoresists," *IEEE Transactions on Electron Devices*, vol. 22, no. 7, pp. 456–464, 1975.

[6] W. D. Hinsberg, F. A. Houle, M. Sanchez, and G. M. Wallraff, "Chemical and physical aspects of the post-exposure baking process used for positive-tone chemically amplified resists," *IBM Journal of Research and Development*, vol. 45, no. 5, pp. 667–682, 2001, doi: 10.1147/rd.455.0667.

[7] H. Ito, "Chemical amplification resists: Inception, implementation in device manufacture, and new developments," *J. Polym. Sci., Part A: Polym. Chem.*, vol. 41, pp. 3863–3870, 2003, doi: 10.1002/pola.10963.

[8] T.-L. Li, "Simulation of the postexposure bake process of chemically amplified resists by reaction-diffusion equations," *Journal of Computational Physics*, vol. 173, no. 1, pp. 348–363, 2001, doi: 10.1006/jcph.2001.6881.

[9] C. A. Mack, "Development of positive photoresists," *Journal of The Electrochemical Society*, vol. 134, no. 1, pp. 148–152, 1987, doi: 10.1149/1.2100396.

[10] C. A. Mack, "New kinetic model for resist dissolution," *Journal of The Electrochemical Society*, vol. 139, no. 8, pp. L35–L37, 1992, doi: 10.1149/1.2221246.

[11] Y. Granik, "Fast pixel-based mask optimization for inverse lithography," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 5, no. 4, Art. no. 043002, 2006, doi: 10.1117/1.2399537.

[12] Y. Kwon and Y. Shin, "Optical proximity correction using bidirectional recurrent neural network," *IEEE Transactions on Semiconductor Manufacturing*, vol. 34, no. 2, pp. 168–176, 2021.

[13] P. Yuan, P. Xu, L. Ma, and Y. Wei, "Optical proximity correction by using unsupervised learning and the patch loss function," *Applied Optics*, vol. 61, no. 14, pp. 3924–3933, 2022, doi: 10.1364/AO.454357.

[14] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019, doi: 10.1016/j.jcp.2018.10.045.

[15] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators," *Nature Machine Intelligence*, vol. 3, no. 3, pp. 218–229, 2021, doi: 10.1038/s42256-021-00302-5.

[16] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Neural operator: Learning maps between function spaces," *Journal of Machine Learning Research*, vol. 24, no. 89, pp. 1–97, 2023.

[17] W. Ye, M. B. Alawieh, Y. Watanabe, S. Nojima, Y. Lin, and D. Z. Pan, "TEMPO: Fast mask topography effect modeling with deep learning," in *Proceedings of the 2020 International Symposium on Physical Design (ISPD)*, 2020, pp. 127–134, doi: 10.1145/3372780.3375565.

[18] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Fourier neural operator for parametric partial differential equations," *arXiv preprint* arXiv:2010.08895, 2020.

[19] Y. Watanabe, T. Kimura, T. Matsunawa, and S. Nojima, "Accurate lithography simulation model based on convolutional neural networks," in *Proceedings of SPIE*, vol. 10454, paper 104540I, 2017, doi: 10.1117/12.2279780.

[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, arXiv:1412.6980.