

Printability Prediction in Projection Two-Photon Lithography via Machine Learning Based Surrogate Modeling of Photopolymerization

Rushil Pingali

G.W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
e-mail: rushilpingali@gatech.edu

Sourabh K. Saha¹

G.W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
e-mail: ssaha8@gatech.edu
ASME Member

Abstract

Two-photon lithography (TPL) is a direct laser writing process that enables the fabrication of cm-scale complex three-dimensional polymeric structures with sub-micrometer resolution. In contrast to the slow and serial writing scheme of conventional TPL, projection TPL (P-TPL) enables rapid printing of entire layers at once. However, process prediction remains a significant challenge in P-TPL due to the lack of computationally efficient models. In this work, we present machine learning-based surrogate models to predict the outcomes of P-TPL to >98% of the accuracy of a physics-based reaction-diffusion finite element simulation. A classification neural network was trained using data generated from the physics-based simulations. This enabled us to achieve computationally efficient and accurate prediction of whether a set of printing conditions will result in precise and controllable polymerization and the desired printing versus no printing or runaway polymerization. We interrogate this surrogate model to investigate the parameter regimes that are promising for successful printing. We predict combinations of photoresist reaction rate constants that are necessary to print for a given set of processing conditions, thereby generating a set of printability maps. The surrogate models reduced the computational time that is required to generate these maps from more than 10 months to less than a second. Thus, these models can enable rapid and informed selection of photoresists and printing parameters during process control and optimization.

Keywords: multi-photon polymerization, additive manufacturing, nanomanufacturing, surrogate modeling, direct laser writing

¹Corresponding author

1. Introduction

Two-photon lithography (TPL) is at the forefront of nanofabrication technologies under development today, promising true three-dimensional (3D) fabrication capabilities with feature resolution below 100 nm [1-3]. TPL systems typically achieve such sub-diffraction-limit resolutions by focusing an intense femtosecond laser beam into a small spot within a curable photopolymer material [4]. Focusing of intense femtosecond light enables one to achieve nonlinear two-photon absorption wherein appreciable light absorption and polymerization-based curing occurs in only a small region within the focal volume [5]. This powerful capability of TPL has been widely applied to fabricate a variety of functional nano-enabled devices, such as microelectronics [6, 7], microfluidics [8, 9], micro-and nano-photonics [10-13], bio-scaffolds and grafts [14, 15], micro-robots [16, 17], and mechanical metamaterials [18, 19].

Given the resolution and versatility of TPL, there have been substantial efforts to scale up the method in terms of throughput, while also reducing cost per unit mass [20-25]. Conventional TPL is slow due to its point-by-point serial writing scheme. One approach to parallelization-based scale up, demonstrated by our group, uses a layer-by-layer projection-based TPL (P-TPL) writing scheme to effectively overcome the traditional rate-versus-resolution tradeoff. In our past work, we have demonstrated P-TPL printing of nanowires with widths and heights smaller than 175 nm and at throughputs 1000 times higher than serial techniques [20]. We have also demonstrated that it is possible to computationally model the P-TPL process using a finite element (FE) based reaction-diffusion simulation that is capable of handling the time and length scales of P-TPL [26].

For further development of P-TPL, it is crucial to not only advance the printing technology to improve performance and reduce cost, but to also accurately characterize the parameter space associated with the printing process. In the absence of reliable characterization data and robust modeling tools, one must rely on iterative trial-and-error printing runs to fabricate the desired features. This imposes a significant barrier to the adoption of the technology for scalable manufacturing of nano-enabled devices. In addition, identifying the process limits is impractical through an iterative empirical approach. For example, choosing the printing conditions and photopolymer material that will produce the highest resolution (i.e., generate the smallest features) is challenging without an accurate model of the process. This is incompatible with a vision of TPL as a mature, scaled-up technology. Computational modeling of the process offers a path towards process scalability.

While we have previously demonstrated the ability to predict the feature size using FE simulations [26], in the present work we approach the modeling problem at a higher level. Here, our aim is to rapidly characterize printability over broad parameter spaces instead of predicting the printed geometry under a specific set of conditions. Physics-based FE simulation is an appropriate choice when prediction of the printed geometry is desired for a specific combination of printing setup and a photoresist material. Under such a specific set of operating conditions, several model parameters can be empirically calibrated to generate an accurate model over a narrow parameter space. However, it is impractical to use FE simulations to model broad parameter spaces wherein multiple parameters can widely vary. This is because it takes several minutes to solve an FE analysis problem for P-TPL. For instance, in this study, we are interested in investigating the process outcome corresponding to 10 input parameters with several parameters varying over multiple orders of magnitude. Producing a process map over such a broad range would require tens of thousands of data points and would take thousands of CPU-hours using FE modeling on cluster computing. Even if such a simulation were to be performed once, it is impractical to repeat this simulation every time a new process map is desired. Here, we adopt a surrogate modeling approach to solve this challenge. By training a classification neural network (CNN) on data points generated from FE simulations, we can

accurately investigate printability regimes with solve times at least seven orders of magnitude faster than the full FE simulation. The surrogate model is presented in Section 2 and the results of the printability mapping are presented in Section 3.

2. Modeling methods

2.1. Projection TPL

The underlying physics of TPL is the two-photon absorption process, in which two photons are simultaneously absorbed by the irradiated material, exciting it into a resonant state [27]. This is a nonlinear process and high intensities are required to achieve appreciable light absorption in UV curable photoresist materials. Generation of free radicals is observed upon relaxation of the photoinitiator molecules from the excited state. It is these free radicals that initiate the polymerization process in the photoresist material [28]. The relationship between the incident light and the resulting number of free radicals generated can be written as [29]:

$$\Delta[R^*] = \frac{D_p \sigma^{(2)} \Phi}{h^2 v^2} [PI] \quad (1)$$

Here, $\Delta[R^*]$ represents the increase in concentration of the primary radicals R^* , measured in mol/dm³, over the duration of a single laser pulse. The primary radicals refer to those radicals that are formed from the photoinitiator molecules. The parameter D_p represents the optical dosage delivered during a single laser pulse. The optical dosage is the square of the instantaneous light intensity, integrated over the duration of a single pulse. The intensity of light increases linearly with the optical power of the beam. The physical significance of the optical dosage is that it is proportional to the amount of energy absorbed per unit volume of the material in the two-photon absorption mode. The parameter $\sigma^{(2)}$ represents the two-photon cross-section of the photoinitiator molecule and the parameter Φ represents the quantum yield of radical formation, i.e., it quantifies the fraction of excited molecules that generate the radicals. The parameter h represents Planck's constant and v is the frequency of the light that illuminates the photoresist material.

In a serial TPL process, the dosage is delivered to the photoresist material sequentially one point at a time. However, in the P-TPL process, an entire 2D light sheet can be illuminated at once, as shown in Figure 1. A digital micromirror device (DMD) is used to selectively illuminate the desired region, acting as a digital mask. Through the simultaneous spatial and temporal focusing of femtosecond light, the entire 2D image can be brought into focus and a whole 2D layer can be printed at once [20].

2.2. Reaction-diffusion model

Previously, we have described a reaction-diffusion modeling scheme for the P-TPL process [26]. The set of chemical equations that capture the essential polymerization process after photoinitiation per equation (1) can be given as:



Here, Eq. (2) represents inhibition of the primary free radical, R^* , by oxygen dissolved in the photoresist medium [30]. Those free radicals that are not immediately quenched go on to start polymerization chain reactions, modeled in equations (3) and (4) as the addition of monomer molecules to a growing polymer chain. The primary radicals react with the monomer molecules (represented by P here) in the photoresist material to generate the secondary radicals P^* . The secondary radicals continue to react with more monomer molecules to form a growing chain of polymer. The number of secondary radicals is preserved during the growth process as each secondary radical forms a new secondary radical upon incorporation into a polymer chain. The polymer chains keep growing until they are terminated through reaction with an oxygen molecule, per equation (5). Other termination mechanisms have been neglected here because oxygen termination has been shown to be the dominant termination mechanism for TPL [31]. The reaction rate constants k_q , k_p , and k_t correspond to the rates of the quenching, polymerization, and termination reactions respectively.

These reaction rate equations can be rewritten as a system of partial differential equations (PDEs) as:

$$\frac{d}{dt}[R^*] = -k_p[P][R^*] - k_q[O_2][R^*] + D_{R^*} \left(\frac{\partial^2 [R^*]}{\partial x^2} + \frac{\partial^2 [R^*]}{\partial z^2} \right) \quad (6)$$

$$\frac{d}{dt}[P] = -k_p[P][R^*] - k_p[P][P^*] \quad (7)$$

$$\frac{d}{dt}[P^*] = k_p[P][R^*] - k_t[O_2][P^*] \quad (8)$$

$$\frac{d}{dt}[O_2] = -k_q[O_2][R^*] - k_t[O_2][P^*] + D_{O_2} \left(\frac{\partial^2 [O_2]}{\partial x^2} + \frac{\partial^2 [O_2]}{\partial z^2} \right) \quad (9)$$

$$\frac{d}{dt}[R^x] = k_q[O_2][R^*] \quad (10)$$

$$\frac{d}{dt}[P^x] = k_t[O_2][P^*] \quad (11)$$

This system of PDE represents the rate of change of the concentration of the reacting chemical species. In addition to changes in concentration caused by the aforementioned chemical reactions, diffusion of O_2 and R^* from regions of high concentration to regions of low concentration is also captured through the spatial derivatives in equations (6) and (9). The slower diffusion of other larger molecules is neglected. Thus, Eqs. (6)-(11) govern the reaction-diffusion mechanism underlying photopolymerization in P-TPL. The initial conditions and the parameters for the simulation are provided in Table 1. For these initial conditions, the properties of pentaerythritol triacrylate (PETA) monomer were used here because PETA has been widely used for TPL printing in the past [31, 32]. It is noteworthy here that despite these fixed initial conditions, a large set of different photoresist material properties could be simulated here because the reaction rate constants, the concentration of the photoinitiator, and the quantum yield of the photoinitiator were allowed to vary over large ranges.

The reaction-diffusion finite element simulation was performed with the commercially available COMSOL Multiphysics software package. The PDE system was solved over a given geometry and for a specified amount of time using a direct time-dependent solver. All simulation in this work was performed for a 2D geometry representing the centermost XZ plane when printing a set of nanowires. It is reasonable to use 2D geometry to model the formation of the nanowires due to the inherent symmetry of system along the length of the nanowires. The simulation has been updated from the version described in more detail previously [26], notable changes are discussed below.

The mesh that the equations were solved over has been updated to the one shown in Figure 2. Instead of a mesh with uniformly-sized elements, this mesh uses progressively finer elements near the focal plane. This mesh has been found to increase the accuracy of printing prediction. Mesh convergence was verified by successively refining the mesh; mesh with values of $n = 40$ were found to be sufficiently fine for simulation of nanowires under most parameter conditions. For challenging parameter regimes (such as at extremely high doses or extremely rapid reactions), a mesh of $n = 50$ was used.

In our prior work [26], the value of the rate constant k_p was modeled as a function of the degree polymer of conversion (DOC) based on literature data, with the other rate constants being fixed. In this work, both the k_p and k_t rate constants were modeled as exponential functions of DOC , with the shape of the function given by parameters A and B , respectively. These functions are:

$$k_p = \left(A - \frac{A^{DOC}}{(A-1)^{DOC-1}} \right) k_{p-0} \quad (12)$$

$$k_t = \left(B - \frac{B^{DOC}}{(B-1)^{DOC-1}} \right) k_{t-0} \quad (13)$$

where k_p and k_t are real numbers. This allows the rate constants to decrease from their initial value to zero as polymerization progresses and the viscosity of the photoresist increases. Similarly, the diffusivity of oxygen is also modeled to decrease linearly with increasing DOC as the reaction progresses. The diffusivity of R^* is approximated as a constant since this species is consumed rapidly at the start of the polymerization.

While the previous version of the simulation was initialized by evaluating the dosage as the product of the square of time-averaged light intensity and a constant pulse duration [26], this version uses dosage values calculated by performing a time-integral of the square of the instantaneous intensity, with integration performed over the duration of the entire pulse. The intensity and pulse durations were evaluated through an optics simulation similar to the one described in the supporting materials of Saha et al. [20]. Since P-TPL uses temporal focusing, dosage values that account for the different pulse durations at different locations are more accurate than dosage values calculated by assuming the same pulse duration at all locations. The model was also updated to account for the consumption of the photoinitiator after each pulse, thereby increasing the accuracy of predictions. Furthermore, the optical parameters (such as magnification and laser repetition rates) were changed to reflect our new P-TPL printing system located at Georgia Tech.

The output of the FE reaction-diffusion simulation is the spatio-temporal evolution of the concentration of each chemical species. From these data, the DOC at the final time is calculated. The DOC refers to the fraction of the monomer that has converted to the polymer; therefore, it characterizes the progress of the polymerization process. From this simulated data, the regions of the photoresist that have cured can be

predicted by applying a threshold on the *DOC*. This thresholding replicates the experimental observation that the photoresist becomes insoluble in solvents when the *DOC* exceeds the threshold *DOC* (DOC_{th}). Regions with *DOC* above this threshold are retained intact after development, i.e., after cleaning the photoresist with solvents. Thus, by applying a threshold, as shown in Figure 3, the shape and size of the printed regions can be discerned. The width and height are measured from this geometric data.

2.3. Training data generation

To train a neural network capable of acting as a surrogate model for the FE simulation, it was necessary to first generate a large amount of training data. This was accomplished through the use of the PACE High Performance Computing (HPC) resource at Georgia Tech. Solving a single model takes on the order of ~ 5 minutes on a workstation computer with an Intel® i7-9700K CPU and 64 GB of RAM; this time can be more or less depending on the simulation parameters. While using HPC does not necessarily speed up the runtime of a single model, it does allow for parallel solving of multiple models simultaneously. This vastly reduces the time needed to generate a large dataset. After some optimization, it was possible to perform on the order of 100 FE model solves per hour, depending on the amount of compute resources allocated.

All simulations for this work were performed with the same projection comprising a set of 5 lines, with a width of 5 pixel and a period of 15 pixel on the DMD. Chemical parameters as well as the duration of exposure and the optical power were varied to capture a wide range of conditions. The 10 input parameters that were varied are listed in Table 2, along with the maximum and minimum values they were allowed to take. The parameters ‘number of pulses’ and the ‘optical power’ represent the illumination-based process parameters that quantify the total amount of optical dosage. The parameters [PI] and Φ represent the process parameters that are determined by the photoinitiator, whereas the rest of the parameters represent the material properties of the monomer in the photoresist. Thus, by varying these three sets of process parameters, one may achieve a vast number of different P-TPL process set ups. It is impractical to attempt the mapping of printability over this entire design space using FE simulations. Instead, samples were generated using FE simulations to train the neural network (NN) models.

A full-factorial sampling scheme was inappropriate for a ten-dimensional parameter space such as this, because aiming to sample each parameter at just 5 levels would require 5^{10} simulations to be performed. This is impractical given the compute time per simulation. Instead, a modified Latin Hypercube Sampling (LHS) scheme was adopted using the `lhsdesign()` function built into MATLAB’s Statistics and Machine Learning Toolbox. The LHS method divides each parameter range into segments and chooses n random points equally distributed among the segments. As a result, it is guaranteed to be more representative of the sample space than pure random selection, such as Monte Carlo sampling [33]. LHS sampling allows for far more dense sampling than factorial sampling, at the cost of not generating all possible combinations of parameter levels.

While the ranges of the rate constants span 8 orders of magnitude, it was thought likely that large portions of this space might not generate any printing. Since there is no benefit to sampling these regions densely, the simulations were performed in batches, with each batch focused on the parameter space deemed to be most promising for printing by the previous batches.

The simulation process was automated through the use of MATLAB LiveLink for COMSOL Multiphysics, which allows algorithmic control of the COMSOL simulation by a MATLAB script. The generation of parameters, batch model solves, and post-processing to find the width and height of the resulting print was automated in this way. In total, 7,084 model solves were completed. Data augmentation

was then performed from these results to increase the size of the dataset that was used for training the neural networks. Due to the nature of the process, it was possible to generate several datapoints from each model solve by varying the DOC_{th} parameter and taking measurements without re-solving the model, since the underlying DOC information remains the same. In addition, it was possible to apply physics-based prior process knowledge to augment the data. For example, it is known that if a given set of parameters failed to print, then decreasing the parameters that are positively correlated with printing (e.g. k_{p-0} , power), while keeping the other parameters constant would also lead to no printing. Here, positive correlation means that individually increasing the input parameter will lead to an increase in the amount of printing. Similarly, increasing the parameters that are negatively correlated with printing (e.g. k_q , k_{t-0}), while keeping the other parameters constant would also lead to no printing. The converse is true for those datapoints that resulted in runaway printing (i.e. over-polymerization). Because the data was created by a deterministic physics-based simulation, it was known with certainty that increasing or decreasing these parameters would always change the outcome in the same direction; that direction can be deduced from the definition of the parameter and the process equations. In this way, synthetic datapoints were generated with high confidence without performing full model solves. The augmented dataset had 26,769 points, i.e., about 4 times the number of models solved. The entire training dataset is available elsewhere [34].

2.4. Neural network surrogate modeling

A surrogate model, or a metamodel, is a computationally inexpensive approximation of a high-fidelity model. Surrogate models are fit to datasets of the type generated here; they are generally not trained on the underlying model physics but solely on input-output relationships. While regression and response surface models have been traditionally popular choices for surrogate models, artificial neural networks (NNs) trained through backpropagation are an increasingly popular choice for high dimensional data [35, 36]. Since this is a high dimensional parameter space with large ranges for each parameter, a NN was chosen over a traditional regression model.

An NN is a machine learning method loosely modeled on the human brain [37, 38]. A network of nodes is set up where each node receives inputs, assigns them weights, and passes them through a transfer function to generate an output. Neural network topologies can include many thousands of layers of nodes. However, for the purpose of this work, a shallow NN with a few layers will suffice [39]. A classification neural net (CNN) applies this principle to classification problems, assigning each input data vector to one of several output classes.

Since a surrogate model inherently has some approximation error, it was not expected that a surrogate model could perform feature size prediction at accuracies comparable to the FE model. Rather, the surrogate model was tasked with predicting whether a given set of printing conditions would result in no printing, printing, or overprinting. These three categories were assigned labels of -1, 0, and 1 respectively. The ground truth labels for the training dataset were generated by analyzing the results of the FE simulations. Prints with connected features or a marked change in aspect ratio were classified as ‘overprinted’. Prints with no regions above DOC_{th} were categorized as having ‘no printing’. Prints with intermediate values of DOC and distinct line features were categorized as ‘printing’. Figure 4 illustrates these categories for a representative value of $DOC_{th}=0.30$. The DOC of 0.30 is not achieved at any spatial location in the ‘no printing’ outcome. In the overprinting case, there is runaway polymerization leading to merging of the line features, i.e., a contiguous patch of material with $DOC>0.30$ exists between the centers of adjacent line features.

Neural network training was performed in MATLAB using the `fitcnet()` function, which allows for the training of a shallow feedforward NN for classification problems. 15% of datapoints were reserved for testing. The rest were used to train a classification neural network with 4 layers of sizes: 15, 9, 5, and 3. ReLU (rectified linear unit) activation functions were used for the hidden layers and a softmax activation function was used for the output layer. All input and output data were treated as real numbers, including the number of pulses (which only took integer values in the training dataset). k -fold cross-validation with $k = 5$ was performed to avoid overfitting. The network was trained on an Intel® i7-1165G7 system with 16GB of RAM. It took an average of 4 minutes to train the net. A schematic representation of the neural network is shown in Figure 5.

3. Results and discussion

3.1. Performance of neural network

The NN was tested on the 15% of reserved datapoints, and it was found to predict the correct result 98.1% of the time. This is possibly a slight overestimation of the true performance, since slight correlations in the dataset might have been introduced during the data augmentation phase. Nonetheless, this performance is generally considered satisfactory for a classification model of this type [39]. The confusion matrix shown in Figure 6 demonstrates that the NN excels at correctly identifying overprinting and non-printing conditions. Accuracy at distinguishing false negatives is slightly lower, with printable conditions being correctly recognized 94.6% of the time. The training dataset contained 39% datapoints of Class -1, 9% datapoints of Class 0, and 52% datapoints of Class 1. As the training dataset consisted of fewer datapoints of Class 0 (printing) in comparison to Class -1 (no printing) and Class 1 (overprinting), so slightly lower accuracy for Class 0 was expected.

Although the accuracy of the NN model is lower than that of the physics-based FE simulations, the NN model has a significant advantage in the computational efficiency. The evaluation time for the NN model is almost instantaneous, taking an average of 0.32 μ s for a single model evaluation. This is significantly shorter than the several minutes taken to solve even a simple 2D FE simulation, with more complex simulations taking even longer. The rapid solve times enable evaluation of tens of thousands of parameter combinations within seconds, accomplishing the main aim of surrogate modeling. The NN model also has an advantage over the FE simulations in the amount of data storage that is required. Each solved FE simulation generates a vast amount of simulation data (\sim 100s of megabyte) that must be parsed to determine the output class corresponding to the printing outcome. In contrast, the NN models directly generate the output class (\sim 1 byte) for a given set of input parameters (\sim 100 byte). Thus, such a surrogate model can be deployed on low-cost computing infrastructure, such as edge computing devices, to perform real-time manufacturing process control and optimization.

3.2. Printability maps

Having generated the surrogate model, it is possible to perform some tests to demonstrate the utility of the model. First, printability mapping was performed by varying two reaction rate constants at a time while holding the other parameters constant. The aim is to verify whether the surrogate model is capable of accurately exploring the material design space, thus enabling choosing reagents with the rate constants required to print under a given set of conditions. To validate this, a random sample of FE simulations was also performed in the same parameter range, to provide a coarse printability map to benchmark against.

The printability map corresponding to varying combinations of polymerization and termination rate constants (i.e., k_{p-0} and k_{t-0}) is shown in Figure 7. All other input process parameters were held constant at a randomly chosen set of values. A DOC_{th} value of 0.20 was used. The background colors represent the

results from the NN, from densely sampled points. The triangles represent FE results for comparison. In this parameter range, it is observed that the NN has captured the essential behavior of the process, albeit with inaccuracies. The NN predicts a gradual shift from non-printing to printing to overprinting as $k_{p,0}$ is increased and $k_{l,0}$ is decreased. This is valid according to the FE model; however, the exact boundaries between the different regimes are not perfectly accurate, as represented by some FE model points being in a different NN region. Nonetheless, the NN has predicted a wide printing band and narrow bands for no printing and overprinting in approximately the correct parameter locations. While an FE model would be necessary to quantify the exact boundaries of printability, the NN can be used to qualitatively assess printability behavior over a broad range of the input parameters.

Similar results were observed when combinations of polymerization and quenching rate constants (i.e., $k_{p,0}$ and k_q) were varied, as illustrated in Figure 8. This printability map shows k_q vs. $k_{p,0}$ with the other input parameters held constant at a different set of values than in Figure 7. Over this parameter space, only printing and no printing outcomes are observed. While the NN results do not exactly match the FE results, the NN is nonetheless able to approximately locate the boundary between the printable and non-printable region. It is interesting to note that despite the inaccuracy in quantitatively locating the boundary, the NN model can qualitatively predict that the boundary gradually shifts toward the top-right corner as the value of $k_{p,0}$ increases. The dotted circles were drawn on Figure 8 to highlight this trend. Thus, the NN surrogate model can be an efficient and effective tool for the prediction of printability over a broad material design space.

A comparison of the compute times required to generate the FE and NN predictions for the two printability maps is summarized in Table 3. Despite the FE sampling being much coarser, generation of the FE datasets takes over an hour, compared to under a second for the NN. Furthermore, the FE solve time depends on how challenging the parameter regime is, which results in an exceedingly long solve time of over 7 hours for the parameter combination in Figure 7. In contrast, the NN solve time does not depend on the input parameter combination. Generation of the printability map of Figure 7 with 100,000 FE simulations would have taken more than 10 months of computation compared to the <1 s taken by the NN model. Thus, the NN surrogate model enables rapid prediction of printability in P-TPL over broad operating regimes, thereby solving a problem that was intractable in the past.

Although our work here has been limited to predicting printability in the P-TPL process, we anticipate that our ML-based approach can also be applied for the printability prediction of other TPL implementations. Our approach is particularly well suited for those TPL implementations that use low repetition rate femtosecond lasers (i.e., rates of \sim 1 kHz), such as the high-throughput implementation demonstrated by Ouyang et al. [40]. Our model of photopolymerization can be adapted to such studies with minimal changes due to the similar time scale of illumination and polymerization reactions. To adapt our work to the more prevalent TPL implementations that use high repetition rate lasers (i.e., rates of \sim 100 MHz), one must modify the polymerization rate equations to include the illumination term within the differential rate equations. Nevertheless, our work demonstrates that ML-based surrogate modeling is an effective and resource efficient approach to predicting printability over a wide range of operating conditions.

As our goal here was to generate ML-based surrogate models of the physics-based FE models, our ML models were trained exclusively with computational data from the FE models. The FE models were themselves validated against experiments [26], but data from experiments were not used to directly train the ML models. In the absence of validated FE models, it is tempting to train the ML models exclusively on experimental data. However, such an approach is more challenging to implement due to the need for a large amount of experimental data and the presence of experimental uncertainty in the data. More

importantly, FE models can generate data over such large operating spaces that cannot be readily accessed through physical experiments with specific systems and photoresist materials. Thus, we anticipate that FE modeling will be an important source of training data for future ML models. Nevertheless, we believe that the accuracy of the ML models can be further improved in the future by fusing the computational dataset with data from physical experiments, as has been demonstrated in ML-based modeling of other AM processes [41].

4. Conclusion

In summary, we have demonstrated that neural network based surrogate modeling is capable of capturing the essential features of the P-TPL parameter space. We have successfully trained a classification model that is capable of evaluating printability for a set of parameters with solve times on the order of microseconds rather than minutes. While the surrogate model currently lacks the accuracy needed for precise feature size prediction, it nonetheless enables computationally cheap qualitative mapping of printability over broad operating regimes. Therefore, it significantly reduces the barriers to the practical adoption of P-TPL as a high-throughput nanoscale additive manufacturing capability by enabling rapid process planning and optimization. Thus, we anticipate that this work will enable scalable nanomanufacturing of complex 3D structures for a variety of devices with applications in computing, energy, transportation, and human health.

Acknowledgments

This work was supported by the National Science Foundation (NSF) CAREER grant 2045147 and in part through research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology. This research was previously presented at the 2023 Manufacturing Science and Engineering Conference (MSEC2023)

Conflict of Interest

The co-authors are inventors on pending and/or issued patent applications on P-TPL for which the intellectual property rights are assigned to either Georgia Tech Research Corporation or the Lawrence Livermore National Security, LLC.

Data Availability Statement

The data sets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

Nomenclature

Chemical Species

PI	photoinitiator
R^*	primary radical
P	monomer
O_2	oxygen
P^*	active polymer chain
R^x	quenched primary radical
P^x	dead polymer chain

Symbols

k_p	polymerization rate constant at a specific degree of polymer conversion
k_{p-0}	polymerization rate constant at the beginning of polymer conversion
k_q	radical quenching rate constant
k_t	termination rate constant at a specific degree of polymer conversion
k_{t-0}	termination rate constant at the beginning of polymer conversion
D_{R^*}	diffusivity of R^*
D_{O_2}	diffusivity of O_2
DOC	degree of polymer conversion
DOC_{th}	threshold degree of polymer conversion
D_p	optical dosage per pulse
$\sigma^{(2)}$	two-photon cross-section of photoinitiator
h	Planck's constant
Φ	quantum yield of photoinitiator
ν	frequency of light
A	k_p exponential function shape parameter
B	k_t exponential function shape parameter

References

- [1] Sun, H.-B., and Kawata, S., 2004, "Two-photon photopolymerization and 3D lithographic microfabrication," *NMR• 3D Analysis• Photopolymerization*, Springer, pp. 169-273.
- [2] Emons, M., Obata, K., Binhammer, T., Ovsianikov, A., Chichkov, B. N., and Morgner, U., 2012, "Two-photon polymerization technique with sub-50 nm resolution by sub-10 fs laser pulses," *Opt. Mater. Express*, 2(7), pp. 942-947.
- [3] LaFratta, C. N., and Baldacchini, T., 2017, "Two-photon polymerization metrology: Characterization methods of mechanisms and microstructures," *Micromachines*, 8(4), p. 101.
- [4] Baldacchini, T., 2015, *Three-dimensional microfabrication using two-photon polymerization: fundamentals, technology, and applications*, William Andrew.
- [5] Tanaka, T., Sun, H.-B., and Kawata, S., 2002, "Rapid sub-diffraction-limit laser micro/nanoprocessing in a threshold material system," *Applied Physics Letters*, 80(2), pp. 312-314.
- [6] Wu, E.-S., Strickler, J. H., Harrell, W. R., and Webb, W. W., "Two-photon lithography for microelectronic application," *SPIE*, pp. 776-782.
- [7] Cumpston, B. H., Ananthavel, S. P., Barlow, S., Dyer, D. L., Ehrlich, J. E., Erskine, L. L., Heikal, A. A., Kuebler, S. M., Lee, I. Y. S., and McCord-Maughon, D., 1999, "Two-photon polymerization initiators for three-dimensional optical data storage and microfabrication," *Nature*, 398(6722), pp. 51-54.
- [8] Vanderpoorten, O., Peter, Q., Challa, P. K., Keyser, U. F., Baumberg, J., Kaminski, C. F., and Knowles, T. P. J., 2019, "Scalable integration of nano-, and microfluidics with hybrid two-photon lithography," *Microsystems & Nanoengineering*, 5(1), p. 40.
- [9] van der Velden, G., Fan, D., and Staufer, U., 2020, "Fabrication of a microfluidic device by using two-photon lithography on a positive photoresist," *Micro and Nano Engineering*, 7, p. 100054.
- [10] Atwater, J., Spinelli, P., Kosten, E., Parsons, J., Van Lare, C., Van de Groep, J., Garcia de Abajo, J., Polman, A., and Atwater, H. J. A. P. L., 2011, "Microphotonic parabolic light directors fabricated by two-photon lithography," 99(15), p. 151113.
- [11] Tian, Y., Kwon, H., Shin, Y. C., and King, G. B., 2014, "Fabrication and characterization of photonic crystals in photopolymer sz2080 by Two-Photon Polymerization using a femtosecond laser," *Journal of Micro and Nano-Manufacturing*, 2(3).
- [12] Moughames, J., Porte, X., Thiel, M., Ulliac, G., Larger, L., Jacquot, M., Kadic, M., and Brunner, D., 2020, "Three-dimensional waveguide interconnects for scalable integration of photonic neural networks," *Optica*, 7(6), pp. 640-646.
- [13] Dietrich, P. I., Blaicher, M., Reuter, I., Billah, M., Hoose, T., Hofmann, A., Caer, C., Dangel, R., Offrein, B., and Troppenz, U., 2018, "In situ 3D nanoprinting of free-form coupling elements for hybrid photonic integration," *Nature Photonics*, 12(4), pp. 241-247.
- [14] Selimis, A., Mironov, V., and Farsari, M., 2015, "Direct laser writing: Principles and materials for scaffold 3D printing," *Microelectronic Engineering*, 132, pp. 83-89.

[15] Limongi, T., Brigo, L., Tirinato, L., Pagliari, F., Gandin, A., Contessotto, P., Giugni, A., and Brusatin, G., 2021, "Three-dimensionally two-photon lithography realized vascular grafts," *Biomedical Materials*, 16(3), p. 035013.

[16] Soreni-Harari, M., St. Pierre, R., McCue, C., Moreno, K., and Bergbreiter, S., 2020, "Multimaterial 3D Printing for Microrobotic Mechanisms," *Soft Robotics*, 7(1), pp. 59-67.

[17] De Marco, C., Alcântara, C. C. J., Kim, S., Briatico, F., Kadioglu, A., de Bernardis, G., Chen, X., Marano, C., Nelson, B. J., and Pané, S., 2019, "Indirect 3D and 4D printing of soft robotic microstructures," *Advanced Materials Technologies*, 4(9), p. 1900332.

[18] Meza, L. R., Das, S., and Greer, J. R., 2014, "Strong, lightweight, and recoverable three-dimensional ceramic nanolattices," *Science*, 345(6202), pp. 1322-1326.

[19] Bauer, J., Meza, L. R., Schaedler, T. A., Schwaiger, R., Zheng, X., and Valdevit, L., 2017, "Nanolattices: an emerging class of mechanical metamaterials," *Advanced Materials*, 29(40), p. 1701850.

[20] Saha, S. K., Wang, D., Nguyen, V. H., Chang, Y., Oakdale, J. S., and Chen, S.-C., 2019, "Scalable submicrometer additive manufacturing," *Science*, 366(6461), pp. 105-109.

[21] Hahn, V., Kiefer, P., Frenzel, T., Qu, J., Blasco, E., Barner-Kowollik, C., and Wegener, M., 2020, "Rapid Assembly of Small Materials Building Blocks (Voxels) into Large Functional 3D Metamaterials," *Advanced Functional Materials*, p. 1907795.

[22] Somers, P., Liang, Z., Johnson, J. E., Boudouris, B. W., Pan, L., and Xu, X., 2021, "Rapid, continuous projection multi-photon 3D printing enabled by spatiotemporal focusing of femtosecond pulses," *Light: Science & Applications*, 10(1), p. 199.

[23] Yang, L., El-Tamer, A., Hinze, U., Li, J., Hu, Y., Huang, W., Chu, J., and Chichkov, B. N., 2015, "Parallel direct laser writing of micro-optical and photonic structures using spatial light modulator," *Optics and Lasers in Engineering*, 70, pp. 26-32.

[24] Vizsnyiczai, G., Kelemen, L., and Ormos, P., 2014, "Holographic multi-focus 3D two-photon polymerization with real-time calculated holograms," *Optics express*, 22(20), pp. 24217-24223.

[25] Jonušauskas, L., Gailevičius, D., Rekštytė, S., Baldacchini, T., Juodkazis, S., and Malinauskas, M., 2019, "Mesoscale laser 3D printing," *Optics express*, 27(11), pp. 15205-15221.

[26] Pingali, R., and Saha, S. K., 2022, "Reaction-Diffusion Modeling of Photopolymerization During Femtosecond Projection Two-Photon Lithography," *Journal of Manufacturing Science and Engineering*, 144(2).

[27] Andrzejewska, E., 2016, "Free radical photopolymerization of multifunctional monomers," *Three-Dimensional Microfabrication Using Two-Photon Polymerization*, Elsevier, pp. 62-81.

[28] Wu, S., Serbin, J., and Gu, M., 2006, "Two-photon polymerisation for three-dimensional micro-fabrication," *Journal of Photochemistry and Photobiology A: Chemistry*, 181(1), pp. 1-11.

[29] Rumi, M., Ehrlich, J. E., Heikal, A. A., Perry, J. W., Barlow, S., Hu, Z., McCord-Maughon, D., Parker, T. C., Röckel, H., Thayumanavan, S., Marder, S. R., Beljonne, D., and Brédas, J.-L., 2000, "Structure-Property Relationships for Two-Photon Absorbing Chromophores: Bis-Donor

Diphenylpolyene and Bis(styryl)benzene Derivatives," Journal of the American Chemical Society, 122(39), pp. 9500-9510.

[30] Jariwala, A. S., Ding, F., Boddapati, A., Breedveld, V., Grover, M. A., Henderson, C. L., and Rosen, D. W., 2011, "Modeling effects of oxygen inhibition in mask-based stereolithography," Rapid Prototyping Journal.

[31] Mueller, J. B., Fischer, J., Mayer, F., Kadic, M., and Wegener, M., 2014, "Polymerization Kinetics in Three-Dimensional Direct Laser Writing," Advanced Materials, 26(38), pp. 6566-6571.

[32] Carlotti, M., and Mattoli, V., 2019, "Functional Materials for Two-Photon Polymerization in Microfabrication," Small, 15(40), p. e1902687.

[33] Stein, M., 1987, "Large sample properties of simulations using Latin hypercube sampling," Technometrics, 29(2), pp. 143-151.

[34] Pingali, R., and Saha, S., 2023, "Data for Printability Prediction in Projection Two-Photon Lithography via Machine Learning Based Surrogate Modeling of Photopolymerization," Mendeley Data, V1, doi: 10.17632/8z29gzb6rd.1, <https://doi.org/10.17632/8z29gzb6rd.1>.

[35] Simpson, T. W., Poplinski, J. D., Koch, P. N., and Allen, J. K., 2001, "Metamodels for computer-based engineering design: survey and recommendations," Engineering with computers, 17, pp. 129-150.

[36] Cao, J., and Lin, Z., 2015, "Extreme learning machines on high dimensional and large data applications: a survey," Mathematical Problems in Engineering, 2015.

[37] Bishop, C. M., 1994, "Neural networks and their applications," Review of scientific instruments, 65(6), pp. 1803-1832.

[38] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., and Arshad, H., 2018, "State-of-the-art in artificial neural network applications: A survey," Heliyon, 4(11), p. e00938.

[39] Zou, J., Han, Y., and So, S.-S., 2009, "Overview of artificial neural networks," Artificial neural networks: methods and applications, pp. 14-22.

[40] Ouyang, W., Xu, X., Lu, W., Zhao, N., Han, F., and Chen, S.-C., 2023, "Ultrafast 3D nanofabrication via digital holography," Nature Communications, 14(1), p. 1716.

[41] Tian, Q., Guo, S., and Guo, Y., 2020, "A physics-driven deep learning model for process-porosity causal relationship and porosity prediction with interpretability in laser metal deposition," CIRP Annals, 69(1), pp. 205-208.

Figure Captions List

Fig. 1: Schematic of P-TPL. (A) Process overview, (B) System overview. From [20]. Reprinted with permission from AAAS.

Fig. 2: Mesh for reaction-diffusion finite elements simulation

Fig. 3: Degree of conversion thresholding. DOC increases from blue to red (top panel). $DOC > DOC_{th}$ shows printed regions only (bottom panel).

Fig. 4: Printing outcomes corresponding to the three output classes. (A) class -1: no printing, (B) class 0: printing, and (C) class 1: overprinting.

Fig. 5: Neural network architecture

Fig. 6: Confusion matrix for the performance of the neural network over the testing dataset that was not used for model training

Fig. 7: Printability map generated via NN and FE models for k_{t-0} ($\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$) vs k_{p-0} ($\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$)

Fig. 8: Printability map generated via NN and FE models for k_q ($\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$) vs k_{p-0} ($\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$)

Table Captions List

Table 1: Initial conditions for simulations

Table 2: Training dataset parameter ranges

Table 3: Comparison of NN and FE model solve times

Figure 1

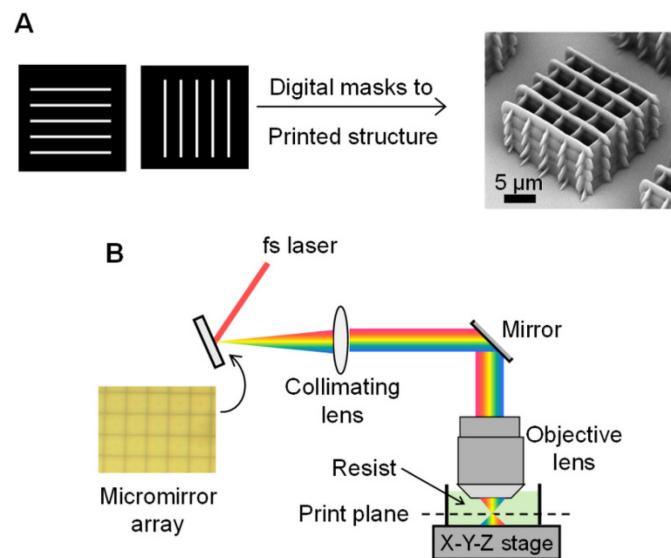


Figure 2

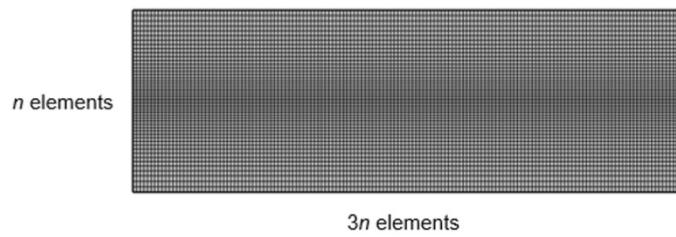


Figure 3

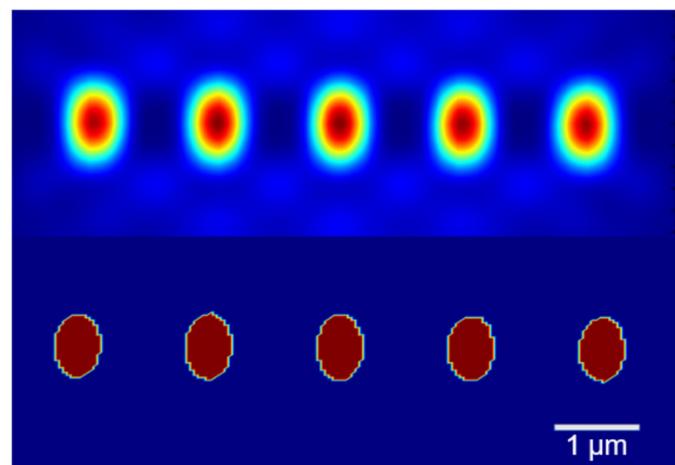


Figure 4

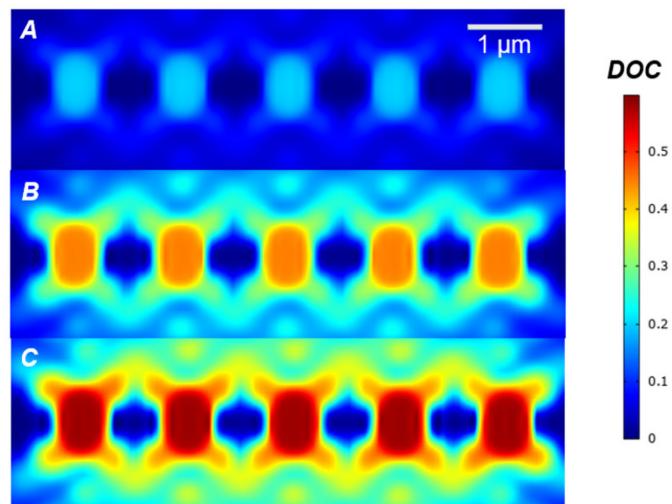


Figure 5

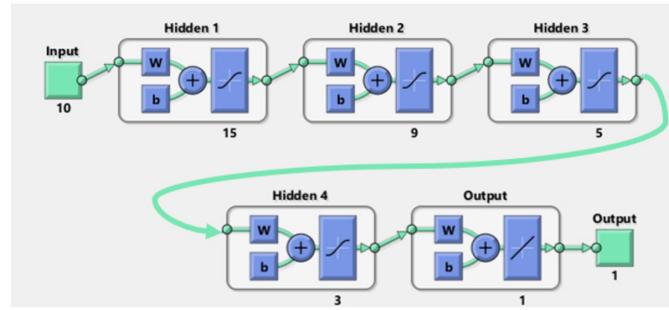


Figure 6

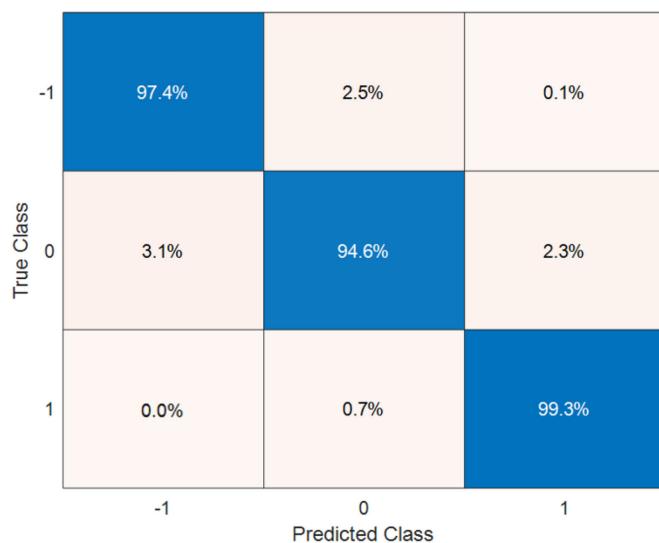


Figure 7

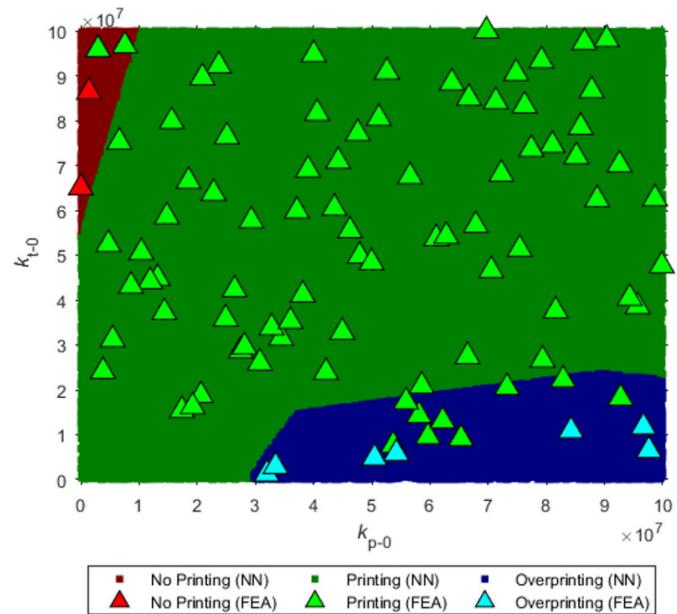


Figure 8

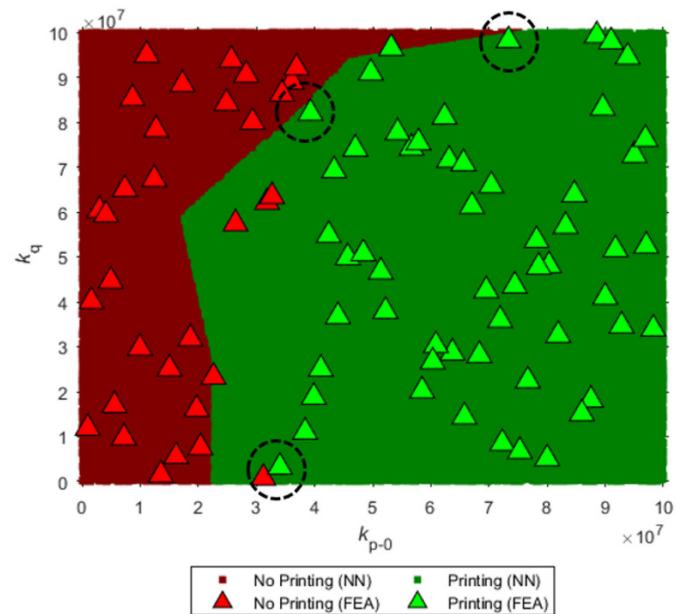


Table 1

Variable	Value	Source
$[R^*]_0, [P^*]_0, [R^x]_0, [P^x]_0$	0 mol dm ⁻³	No light exposure before printing
$[P]_0$	3.956 mol dm ⁻³	Material datasheet
$[O_2]_0$	6×10^{-3} mol dm ⁻³	Mueller et al. [31]
D_{O2}	1.2×10^{-12} m ² s ⁻¹	Estimated with Stokes-Einstein equation
D_{R^*}	10^{-13} m ² s ⁻¹	
Pulse repetition rate	5 kHz	Printing conditions

Table 2

Parameter	Lower Bound	Upper Bound	Units
k_{p-0}	1	1×10^8	$\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$
k_q	1	1×10^8	$\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$
k_{t-0}	1	1×10^8	$\text{dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$
Number of pulses	1	20	-
A	1×10^{-20}	1×10^{-10}	-
B	1×10^{-20}	1×10^{-10}	-
Optical power	0.01	0.50	W
Φ	1×10^{-4}	1×10^{-2}	-
[PI]	1×10^{-4}	1×10^{-2}	mol/dm^3
DOC_{th}	0.02	0.70	-

Table 3

Parameter combinations	Number of datapoints	Time taken to generate
Figure 7: NN	100,000	<1s
Figure 7: FE	96	7h 14min
Figure 8: NN	100,000	<1s
Figure 8: FE	96	1h 19min