



# Mass Transport in Confined Micro Geometry in Reciprocating Paddle Plating Cell by Numerical Simulation and Neural Network Analysis

Q. Huang<sup>\*,z</sup> 

Department of Chemical and Biological Engineering, The University of Alabama, Tuscaloosa, Alabama 35405, United States of America

The mass transport in confined geometry in a paddle plating cell is studied using numerical simulation with an attempt to extrapolate an explicit correlation to understand the transport physics and to predict the electrodeposition rate of metal microstructures. A moving boundary in conjunction with mapped mesh is used to allow the reciprocating movement of the flow. A correlation is obtained based on a generalized additive model using multivariate linear regression. Neural networks are also used to analyze the efficacy of such correlation and to determine the descriptor characterizing the error in prediction. A two-step convective mass transport process, one in the bulk electrolyte outside the patterns and the other inside the micro-trenches, is demonstrated to better describe the overall transport physics and improve the correlation.

© 2024 The Electrochemical Society ("ECS"). Published on behalf of ECS by IOP Publishing Limited. All rights, including for text and data mining, AI training, and similar technologies, are reserved. [DOI: [10.1149/1945-7111/ad7c83](https://doi.org/10.1149/1945-7111/ad7c83)]

Manuscript submitted July 2, 2024; revised manuscript received September 6, 2024. Published October 7, 2024.

Supplementary material for this article is available [online](#)

Electroplated copper micro bumps have been used ubiquitously in the 3D integration and packaging of the state-of-the-art semiconductor devices.<sup>1–3</sup> As the technology continues to advance, the micro bumps continue to scale from several hundred microns to tens of microns. This scaling in size requires tighter control of the growth in the confined geometry in terms of not only the total amount of the metal but also the topography and uniformity of the metal structures.<sup>4–6</sup> These requirements are to be achieved without compromising a fast rate of growth.

Various hydrodynamic designs have been adopted to facilitate the mass transport into small geometries, thus expediting the metal growth rate and improving the throughput of this manufacturing process. For example, fountain flow cells, where the electrolyte moves vertically in the normal direction of a rotating wafer surface,<sup>7,8</sup> have been successfully implemented in damascene copper plating. The super-conformal copper plating in the damascene processes is controlled by the synergistic interactions of organic additives and the plating is typically operated at kinetically controlled regime. On the other hand, plating cells with impinging jet flows,<sup>9–12</sup> a reciprocating paddle,<sup>13–15</sup> or a vibrational shear plate<sup>16</sup> are designed to provide effective agitation, thinning the boundary layer and increasing the convective mass transport rate. Among them, paddle cell was initially invented for magnetic alloy thin film plating, where an external stationary magnetic field can be applied during plating to define the magnetic anisotropy in the plated films.<sup>13,17,18</sup> More importantly, the paddle also enables a strong but well-controlled convective mass transport of the metal cations, mitigates the high pH at the substrate surface, and prevents the formation and inclusion of oxides, resulting in high quality magnetic films with uniform and well defined compositions.<sup>19</sup> The paddle cell design has also been adopted for copper bump plating, where the metal growth rate in confined geometry is to be maximized.

The mass transport on flat substrates under the reciprocating paddle agitation has been studied both experimentally and numerically. For example, an early experimental study of the mass transfer using stripe-shaped electrodes shows that the current experiences a periodic modulation depending on the instantaneous position of the paddle relative to the electrode. A generalized mass transfer rate constant is obtained solely dependent on the Reynolds number and Schmidt number.<sup>14,20</sup> On the

other hand, numerical simulation of a moving paddle showed that the a bluff body flow is created due to the generation and shedding of the vortices.<sup>21</sup> In this case, the Strouhal number describing the vortex shedding frequency needs to be acquired from the numerical simulation and be used to describe the mass transport rate. However, all these prior studies were based on flat substrates. The question remains how such paddle flow impacts the mass transport when the deposition is confined in small and recessed regions in the substrate. To the best of our knowledge, such studies have not been reported and the question remains unanswered. This paper reports a systematic study on this subject using numerical simulation and data analysis.

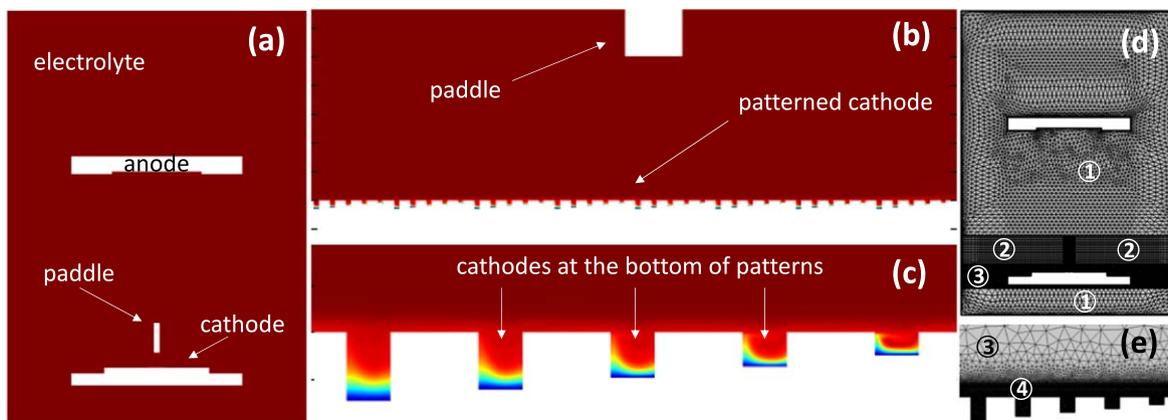
## Numerical Simulation

**Physics model.**—In this study, 2-dimensional models are used to describe and simulate the paddle cell. While the patterns in 2D models represent trenches and 3D models are required to simulate the via structures in reality, the 3D models would exponentially increase the degree of freedom in computation, rendering the computation time much longer and the convergence almost impossible to reach. Figures 1a to 1c show the diagram of the paddle cell. The detailed geometry and typical dimensions are presented in Fig. S-1 in Supplementary Information. It consists of the anode and cathode facing each other and a rectangular paddle placed at a fixed distance from the cathode surface. The anode, patterned cathode, and paddle are all aligned with the center of the cell. The anode is a flat metal surface and the cathode consists of 40 trenches grouped into 8 identical sets and each set consists of 5 patterns with a same width and five different depth. As discussed later, this variety of depth in a single model increases the number of cases for analysis without significantly compromising the result accuracy. The pitch (center-center distance) between two adjacent trenches is fixed as 3 times the feature width, which is typical for microbump patterns.

The model consists of momentum and mass transfer. The latter includes both diffusive and convective transfer. The electromigration contribution for the mass transfer is ignored in this study as a highly conductive electrolyte is typically used in micro-bump electroplating. The electrolyte is Newtonian and incompressible with uniform viscosity. The gravitational force is ignored. In addition, the local changes in electrolyte density and ionic diffusivity resulted from local concentration variation are expected to be small<sup>22–25</sup> and are ignored. The governing equations in the electrolyte include

\*Electrochemical Society Member.

<sup>z</sup>E-mail: [qhuang@eng.ua.edu](mailto:qhuang@eng.ua.edu)



**Figure 1.** Schematic diagrams of the (a) overall paddle cell model; (b) magnified area between paddle and patterned cathode with 40 trenches; (c) one of the eight groups of five trenches with a typical concentration profile upon paddle agitation; (d) the meshing of the entire paddle cell; and (e) refined meshing in the area of patterned cathode.

Continuity:

$$\rho \nabla \cdot u = 0 \quad [1]$$

Momentum conservation:

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = -\frac{1}{\rho} \nabla p + \nu \cdot \nabla^2 u \quad [2]$$

Mass conservation:

$$\frac{\partial C_i}{\partial t} + \nabla \cdot (-D_i \nabla C_i + C_i u) = 0 \quad [3]$$

where  $t$ ,  $\rho$ ,  $u$ ,  $p$ ,  $\nu$ ,  $D_i$ ,  $C_i$  represent time, electrolyte density, velocity, pressure, kinematic viscosity, and the diffusion coefficient and concentration of the species of interest, i.e.,  $\text{Cu}^{2+}$  in this study. Typical boundary conditions for fluidic dynamics are applied for most boundaries except for the paddle. The reciprocating movement of paddle is achieved using a moving mesh method. The mesh in the regions ② in Fig. 1d deforms laterally according to the position of paddle. The position of paddle follows a Sinusoidal function when a rotary motor is used to drive the paddle,<sup>19</sup>

$$x(t) = \frac{s}{2} \cdot \sin(2\pi f \cdot t) \quad [4]$$

where  $s$ ,  $f$ ,  $t$  are the stroke length of paddle, frequency of paddle, and time, respectively. While a linear motor can also be used to drive a paddle cell, the sinusoidal function is used here as it naturally avoids singularity and facilitates model convergence. On the other hand, because the total length of the patterned cathode accounts for a small portion of the stroke, the velocity at which the paddle sweeps on top of the cathode pattern can be approximated as a constant, i.e., the maximum speed at the central position

$$u(t) = \pi f \cdot s \cdot \cos(2\pi f \cdot t) \quad [5]$$

$$u_{\max} = \pi \cdot s \cdot f \quad [6]$$

The paddle is stationary before the start of simulation, or  $t < 0$ , and the velocity at  $t = 0$  is supposed to be  $u_{\max}$ , a non-zero finite value, resulting in a step function and complicating the numerical computation. This step function is smoothed out across the first 1/4 paddle cycle to improve the computation.

Most of boundary in the mass transfer physics are set as no flux condition. The concentration at the anode surface is fixed as the bulk concentration. In this study, the concentration at cathode surface, namely at the bottom of the trench patterns, is set as 0. No electrochemistry such as Butler-Volmer or Tafel conditions are

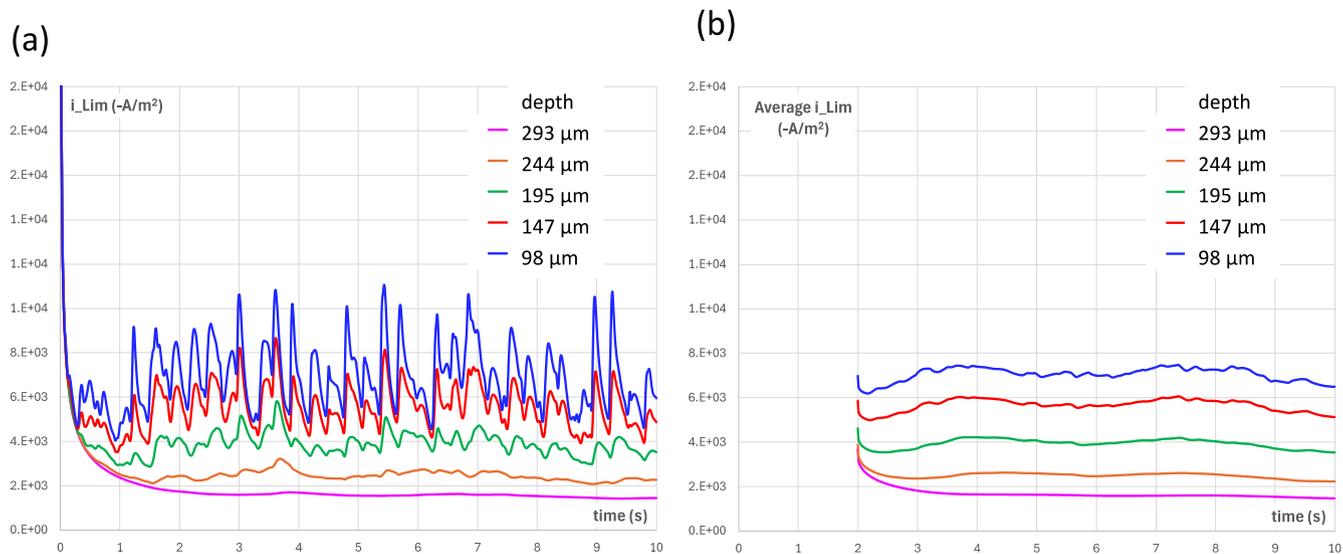
used. The initial concentration in the electrolyte is set as bulk concentration. However, a much lower initial concentration inside the patterns is found to facilitate computation by avoiding a big flux at  $t = 0$ . The limiting current density is calculated from the flux at the cathode surface using Faraday's law

$$i_{\text{lim}} = n \cdot F \cdot N_{\text{Cu}^{2+}} \quad [7]$$

where  $n$  is the number of electrons transferred per copper ion,  $F$  is the Faraday's constant, and  $N_{\text{Cu}^{2+}}$  stands for the flux of copper ions obtained from the simulation. As shown in Figs. 1d and 1e, the meshing is performed separately for 4 regions. Other than the mapped mesh in the deforming regions ②, triangular meshes are used for all the other regions. The region around cathode, i.e. region ③, requires finer mesh and the mesh inside the trench patterns, or region ④, is even finer. The boundary layer mesh at the bottom of pattern needs to be thin enough to allow the establishment of the concentration gradient at the beginning of the computation. As mentioned above, the fluidic properties of the electrolyte, such as the density and viscosity, are approximated as constants independent from the concentration. The fluid mechanics and the mass transfer physics can be solved separately, where the velocity profile from the former is solved first and used as the input for the latter. Compared with solving the two physics simultaneously, this sequential approach improves the convergence and speeds up the computation.

**Electrolyte properties.**—Highly concentrate  $\text{CuSO}_4$  electrolytes are typically used in microbump fabrication to enable a fast plating rate. A high concentration of  $\text{H}_2\text{SO}_4$  is also preferred to enhance the conductivity of electrolyte and enable high plating current. However, the solubility of  $\text{CuSO}_4$  is limited in the presence of a high concentration of  $\text{H}_2\text{SO}_4$ . A moderately higher temperature can be used in manufacturing not only to enable a higher metal concentration and a high conductivity but also to enhance the mass transfer and deposition kinetics. The physical properties of two typical solutions are listed in Table I. The diffusion coefficients of  $\text{Cu}^{2+}$  in the example electrolytes are determined using limiting current measurements on a rotating disc electrode at different rotation rates. A wide range covering these measured values are used to randomize the parameters for the simulation.

**Computation flow.**—All numerical simulations in this study are performed using COMSOL Multiphysics software. A flow chart is provided in Fig. S-2 in Supplementary Information detailing the procedure of COMSOL simulation. In brief, the simulation of the non-steady state physics is carried out for a physical time of 10 s with a data acquisition time step of 0.01 s. The entire COMSOL physics model is parameterized, including the cell geometry, pattern geometry, paddle movement, and electrolyte properties. An external



**Figure 2.** (a) The time transient of the overall limiting current densities for patterned features with different depth, and (b) the limiting current densities averaged across the 2 s time frame preceding the time snap.

**Table I. Density, viscosity, and copper ion diffusivity in typical electrolytes and the range of values used in the simulation.**

	Electrolyte#1	Electrolyte#2	Range used
CuSO <sub>4</sub> (mol/m <sup>3</sup> )	630	787	[500, 1000]
H <sub>2</sub> SO <sub>4</sub> (M)	1.4	0.25	
Temperature (C)	40	25	
Density (kg/m <sup>3</sup> )	1169	1133	[800, 1600]
Dynamic viscosity (Pais·s)	$1.20 \times 10^{-3}$	$1.49 \times 10^{-3}$	
Kinematic viscosity (m <sup>2</sup> /s)	$1.02 \times 10^{-6}$	$1.32 \times 10^{-6}$	$[0.8, 1.6] \times 10^{-6}$
Cu <sup>2+</sup> Diffusion coefficients (m <sup>2</sup> /s)	$7.80 \times 10^{-10}$	$8.17 \times 10^{-10}$	$[2, 20] \times 10^{-10}$

Matlab program randomizes these parameters, uses each set of parameters to call the model through a Matlab-COMSOL interface, and exports the limiting current density for each cathode pattern. In certain studies, where the impact of a specific parameter is studied, the parameters are manually set in the external Matlab code. A windows server with Intel Xeon Gold 5220R CPU and 96 MB error correction RAM is used for the simulation. Depending on the trench dimensions and paddle movement in each case, it takes between an hour to a few days to complete the 10 s simulation.

## Results and Discussion

**COMSOL simulation.**—The typical evolution of the velocity and concentration profiles in the cell are shown in the GIF animations in Supplementary Information. The animation titled as “Animation velocity” shows the overall transient velocity profile across the entire cell, “Animation velocity zoomed” shows the velocity profile at one of the eight sets of trenches, and “Animation concentration” shows the concentration profile at the same set of the trenches with more details explained in Fig. 1c. The concentration varies only in the vicinity of the cathode and stays as the bulk concentration across the other regions of the cell. Figure 2a shows the non-steady state transient of the limiting current density. It is extremely high at  $t = 0$  and decays rapidly as the concentration gradient is being established. The current can reach a minimal value after one or two seconds because the concentration gradient develops quickly and the fluid flow pattern is yet to be

established across the cell. The limiting current density continues to fluctuate as the paddle sweeps back and forth on the cathode. For this particular case in Fig. 2, the paddle moves at a frequency of 1.68 Hz, or 101 strokes per minute. The width of the five cathode patterns is 187  $\mu\text{m}$ , and the depth are 293, 244, 195, 147 and 98  $\mu\text{m}$  respectively. As expected, the current fluctuation is much more pronounced for shallower patterns. While the current fluctuation is not identical for every stroke, approximately 10 spikes are observed across 3 s, corresponding to the fact that paddle sweeps across the cathode twice per stroke, thus 3.36 sweeps per second. Figure 2b shows the limiting current density averaged across a period of 2 s. This time average stabilizes after a few strokes. The average limiting current densities between  $t = 5$  to 10 s are reported in all following studies.

Simulations are first carried out to determine the current distribution and the effect of using a mixture of different trench depths in a single model. Figure S-3 in Supplementary Information compares the limiting current density simulated using two different approaches. The first one uses models with 8 sets of trenches and 5 trenches with different depths in each set (as explained in Fig. 1). The second approach uses 40 identical trenches and is repeated 5 times to simulate 5 different trench depth. First, slightly higher current densities are observed in the features at the end of patterns. This is commonly observed in manufacturing due to the spherical diffusion at pattern corners.<sup>26</sup> Dummy features are often included in pattern design to avoid big difference between patterned and blanket areas. The shallower features in the first approach, i.e., when they are combined with deeper features, have a higher limiting current density. This is believed due to a less depleted concentration at the deeper neighbours enabling extra supply of the species for the shallower neighbours. On the other hand, the deeper features have lower limiting current density when they are next to shallower features. Nevertheless, the difference between the two approaches are less than 10% for all features. As the use of mixed trenches in a single model accelerates the simulation and increases the available data by 5 times, the first approach is used across this study.

The reproducibility of numerical simulation is also verified using 25 different trenches and the results are presented in Fig. S-4 in Supplementary Information. It is evident that the simulation is very reproducible with the variation below 1% for most of cases and below 4% for all cases.

While the numerical simulations in Figs. S-3 and S-4 use preset parameters, a large amount of simulation using randomized parameters are performed to derive a correlational method to predict limiting current densities from plating conditions. In detail, the plating conditions are described using eight parameters, paddle stroke length (s), paddle

frequency ( $f$ ),  $\text{Cu}^{2+}$  concentration ( $C$ ), diffusion coefficient or diffusivity ( $D$ ), electrolyte density ( $\rho$ ), kinematic viscosity ( $\nu$ ), trench width ( $w$ ), and trench depth ( $h$ ). In total, 7000 trenches are simulated using 1400 COMSOL models. Among them, 5000 randomized trenches are simulated first for data fitting. The additional 2000 randomized trenches are simulated separately and are used to evaluate the prediction by the fitted correlation.

**GAM model.**—The simulated limiting current densities are first fitted using a generalized additive model (GAM). In the first attempt, the maximum paddle velocity at central position (in Eq. 6) is used

$$\begin{aligned} \ln(i_{Lim}) = & k_0 + k_u \cdot \ln(u) + k_C \cdot \ln(C) \\ & + k_D \cdot \ln(D) + k_\rho \cdot \ln(\rho) \\ & + k_\nu \cdot \ln(\nu) + k_w \cdot \ln(w) + k_h \cdot \ln(h) \end{aligned} \quad [8]$$

where  $C$ ,  $D$ ,  $\rho$ ,  $\nu$ ,  $w$ ,  $h$  represent the randomized concentration, diffusivity, density, kinematic viscosity, trench width, and trench depth, respectively.  $u = \pi \cdot s \cdot f$  represents the paddle velocity at central position. The  $k$ 's are fitted constants. This is equivalent to a power correlation

$$i_{Lim} = e^{k_0} \cdot u^{k_u} \cdot C^{k_C} \cdot D^{k_D} \cdot \rho^{k_\rho} \cdot \nu^{k_\nu} \cdot w^{k_w} \cdot h^{k_h} \quad [9]$$

Figure 3 shows the results of this GAM model, where  $k_0 = 9.676$ ,  $k_u = 0.442$ ,  $k_C = 0.997$ ,  $k_D = 0.699$ ,  $k_\rho = 0.001$ ,  $k_\nu = -0.193$ ,  $k_w = 0.692$ , and  $k_h = -1.136$ . The mean and maximum absolute percent error (APE) of fitted limiting current for the 5000 cases are 12.8% and 65.1%, respectively. The same model is then used to predict the limiting current for additional 2000 randomized cases, and the mean and maximum APE are 12.7% and 60.1%, almost identical to the fitting. Because these two sets of data are randomized separately, they remain truly independent and the errors are found always similar regardless of methods of fitting.

The fitted value for  $k_C$ , 0.997 is very close to 1, consistent with the expectation that the mass transport flux is proportional to the cation concentration. Furthermore,  $k_\rho$  is 0.001, close to 0, suggesting that the effect of density becomes obsolete in the presence of kinematic viscosity. With the two parameters of concentration and density omitted, the fitted limiting current density follows

$$i_{Lim\_GAM} = e^{9.676} \cdot C \cdot u^{0.442} \cdot D^{0.699} \cdot \nu^{-0.193} \cdot w^{0.692} \cdot h^{-1.136} \quad [10]$$

The mass transport Sherwood number can be defined as

$$Sh = h_{mass} \cdot \frac{L_c}{D} = \frac{i_{Lim}}{nF \cdot C} \cdot \frac{L_c}{D} \quad [11]$$

where  $h_{mass}$  is the overall convective mass transport coefficient and  $L_c$  is the characteristic length to describe the convection behavior. Using the fitted limiting current density  $i_{Lim}$ , the Sherwood number can be described as

$$\begin{aligned} Sh_{GAM} &= 0.0826 \cdot L_c \cdot u^{0.442} \cdot D^{-0.301} \cdot \nu^{-0.193} \cdot w^{0.692} \cdot h^{-1.136} \\ &= 0.0826 \cdot L_c^{0.558} \cdot \left(\frac{u \cdot L_c}{\nu}\right)^{0.442} \cdot \left(\frac{\nu}{D}\right)^{0.301} \cdot (\nu)^{-0.052} \cdot w^{0.692} \cdot h^{-1.136} \end{aligned} \quad [12]$$

The Reynolds and Schmidt numbers can be defined as  $Re = \frac{u \cdot L_c}{\nu}$  and  $Sc = \frac{\nu}{D}$ , respectively. Therefore,

$$Sh = 0.0826 \cdot L_c^{0.558} \cdot (Re)^{0.442} \cdot (Sc)^{0.301} \cdot (\nu)^{-0.052} \cdot w^{0.692} \cdot h^{-1.136} \quad [13]$$

While the additional effect of kinematic viscosity does not completely vanish, it becomes minimal, suggesting its main impact is

included in the Reynolds and Schmidt numbers. Equation 13 is consistent with common convective mass transport processes,<sup>27,28</sup> where the Sherwood is correlated with Reynolds and Schmidt numbers,  $Sh = k \cdot (Re)^m \cdot (Sc)^n$ , where  $k$ ,  $m$ , and  $n$  are constants and  $n$  is typically between 0.3 to 0.4. A similar correlation,

$$Sh \approx 0.0826 \cdot (Re)^{0.442} \cdot (Sc)^{0.301} \quad [14]$$

would be obtained if the combination of  $L_c^{0.558} \cdot w^{0.692} \cdot h^{-1.136}$  is dimensionless, or in other words,  $L_c \propto w^{-1.24} \cdot h^{2.04}$ . However, the latter has a dimension of  $[m]^{0.8}$ . While  $L_c$  and the convective mass transport are dominated by the dimension of trench features, it is also impacted by other dimensions of the system. The power of Schmidt number,  $n$ , is 0.301 here, slightly different from 1/3, which is commonly observed in convective transport.<sup>29</sup> The value of 1/3 has also been reported or used in previous studies on flat surface in a paddle cell.<sup>14,20,21</sup> As discussed later, it will be seen that the convective transport in the confined trenches can be viewed as multiple steps. The combination of external and internal fluidic flow, the time averaging over the periodic paddle flow, and the spherical diffusion at the trench opening are believed to contribute to this slightly different  $n$  of 0.301.

While not included in detailed discussion here, some additional simulation cases are also carried out using randomized dimensions of the paddle. The results show that the mass transport rate increases with the paddle height ( $H$ ). On the other hand, the paddle width ( $L$ ) and the gap distance between paddle and cathode ( $G$ ) have minimal impacts. This is different from some experimental studies on flat surface, where the characteristic length was found to be  $L_c = H + G$ .<sup>20</sup> From the Eqs. 11 to 14, this  $L_c$  suggests the mass transport rate would decrease with the increase of paddle height and gap distance. However, a numerical simulation study on flat surface showed that the mass transport increases with the paddle height,<sup>21</sup> consistent with the observation in this study.

As the limiting current density reported in this study is an value averaged across multiple strokes, the effects of different parameters can be further complicated. In the second attempt, the GAM fitting approach is also performed using the stroke length ( $s$ ) and frequency ( $f$ ) separately instead of a single value of velocity. The following correlation is obtained with an average error of 12.4% and a maximum error of 60.0%, similar to the fitting using the combined velocity.

$$i_{Lim\_GAM} = e^{9.567} \cdot C \cdot s^{0.229} \cdot f^{0.464} \cdot D^{0.701} \cdot \nu^{-0.192} \cdot w^{0.687} \cdot h^{-1.129} \quad [15]$$

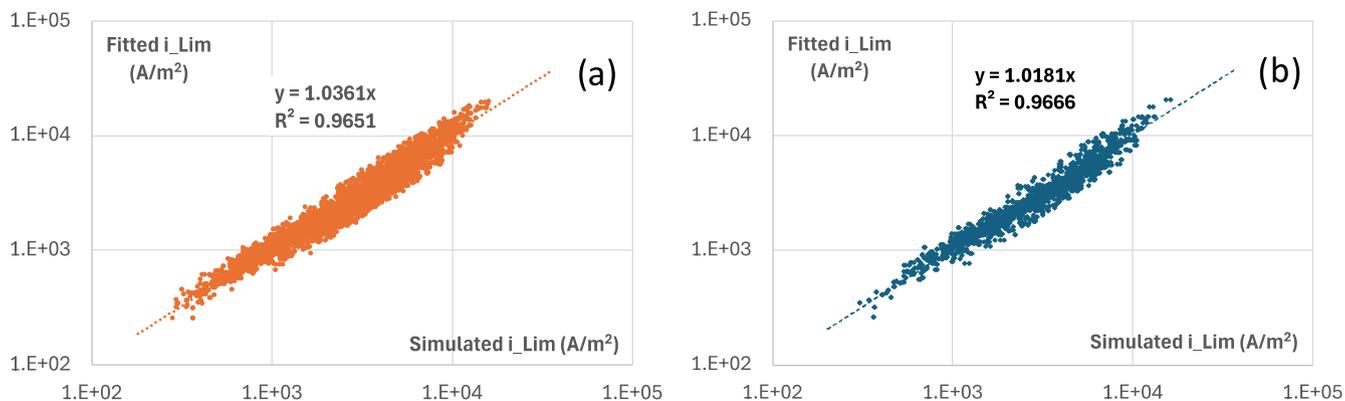
$$Sh_{GAM} = 0.074 \cdot L_c \cdot s^{0.229} \cdot f^{0.464} \cdot D^{-0.299} \cdot \nu^{-0.192} \cdot w^{0.687} \cdot h^{-1.129} \quad [16]$$

The limiting current density in this GAM is similar to the fitting using the combined velocity except that the power coefficients for stroke length and frequency are different. When a combined velocity  $\pi s f$  is used in the Reynolds number, the Sherwood number can be described as

$$Sh \approx 0.0435 \cdot (Re)^{0.464} \cdot (Sc)^{0.299} \quad [17]$$

with the characteristic length  $L_c \propto s^{0.44} \cdot w^{-1.28} \cdot h^{2.11}$ . It has a dimension of  $[m]^{1.27}$ , also suggesting additional contributions from other dimensions in the system. In other words, separating stroke ( $s$ ) and frequency ( $f$ ) does not significantly improve the GAM correlation.

Further analysis of the GAM model is carried out to account for the time averaging effect. As Fig. 2a shows, the typical time evolution of the limiting current density includes spikes of current due to the sweep of paddle followed by current decay as the paddle moves away and the fluid flow in the trench quickly slows down. Assuming the mass transport flux due to a convective flow can be well described using a correlation  $Sh = k \cdot (Re)^m \cdot (Sc)^n$  and the



**Figure 3.** The relationship between numerically simulated and fitted limiting current using the GAM: (a) 5000 cases for fitting and (b) 2000 cases for testing.

velocity decays exponentially following  $u = \pi \cdot s \cdot f \cdot \exp\left(-\frac{t}{\tau}\right)$  with a time constant  $\tau$ , the averaged Sherwood number can be obtained by integrating the transport rate between two sweeps, namely across a time frame  $t \in \left[0, \frac{1}{2f}\right]$ . In this case,

$$\begin{aligned} \overline{Sh} &\propto \frac{1}{\frac{1}{2f}} \cdot \int_0^{\frac{1}{2f}} (\pi s f)^m \cdot \exp\left(-\frac{m \cdot t}{\tau}\right) \cdot dt \\ &= s^m \cdot f^{m+1} \cdot \frac{\tau}{m} \cdot \left[1 - \exp\left(-\frac{m}{2 \cdot f \cdot \tau}\right)\right] \end{aligned} \quad [18]$$

For the two extreme cases,

$$\text{if } \frac{m}{2 \cdot f \cdot \tau} \gg 1, \text{ then, } \overline{Sh} \propto \tau \cdot s^m \cdot f^{m+1} \quad [19]$$

$$\text{if } \frac{m}{2 \cdot f \cdot \tau} \ll 1, \text{ then, } \overline{Sh} \propto s^m \cdot f^m \quad [20]$$

In reality,  $m$  and  $f$  are both on the order of 1. The decay time constant  $\tau$  is dependent on the fluidic property and trench geometry. The animations of fluidic velocity profile are included in the Supplementary Information and two additional snapshots are presented in Fig. 4. The fluidic velocity in bulk solution estimated from the vortex velocity in Fig. 4b decreases from  $0.65 \text{ m s}^{-1}$  to  $0.32 \text{ m s}^{-1}$  (50% decrease) across 0.24 s, suggesting a time constant of about 0.34 s, which is also on the order of 1. On the other hand, the eddy flow in the trenches in Fig. 4a decreases from 8 to  $3 \text{ mm s}^{-1}$  within 0.12 s, corresponding to a time constant of 0.13 s, also on the order of 1. Therefore, for most cases, the integrated average Sherwood number is expected to be in between the two extreme cases in Eqs. 19 and 20. For these intermediate situations,

$$\overline{Sh} \propto \tau^k \cdot s^m \cdot f^r, \quad 0 < k < 1, \quad m < r < m + 1 \quad [21]$$

This is consistent with the GAM fitting in Eq. 16, where the  $\overline{Sh} \propto s^{0.229} \cdot f^{0.464}$ . As far as the time constant of the velocity decay is concerned, it depends on how rapidly the kinetic energy of fluidic flow dissipates. It is inversely correlated with the kinematic viscosity, resulting in an additional small contribution from the kinematic viscosity in Eqs. 12 and 16.

**Neural network.**—A neural network is used to further evaluate the impacts of each parameters in the mass transport rate. The fitnet function in Matlab neural network fitting package is used with the number of neurons in each layer specifically defined. The overall workflow is shown in Fig. 5. The inputs are still the same, stroke length ( $s$ ), frequency ( $f$ ), diffusivity ( $D$ ), density ( $\rho$ ), kinematic viscosity ( $\nu$ ), trench width ( $w$ ), and trench depth ( $h$ ), and the target is

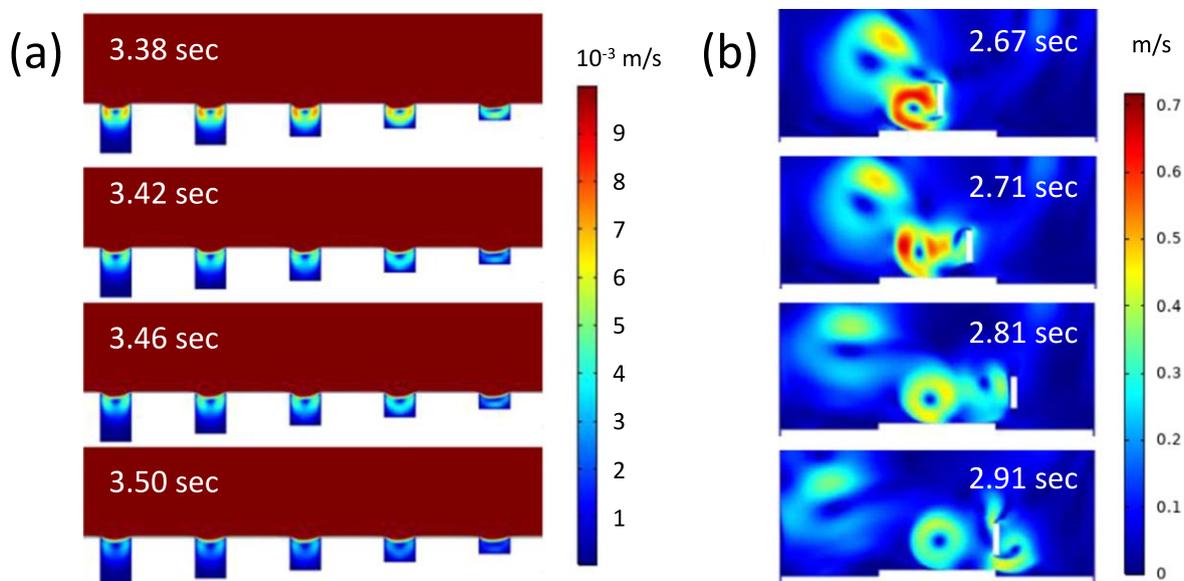
the limiting current density normalized with concentration ( $i_{Lim}/C$ ). The first batch of 5000 COMSOL simulated cases are always used to fit the neural network (NN) and the second batch of 2000 cases are always used for NN prediction and evaluation. Among the first 5000 cases, 80% are used for training, 10% for validation, and another 10% for testing. This randomized 80–10–10 partition is done automatically in the fitnet function in Matlab. While the fitnet function provides an error during the testing step using 10% of data, the second batch of 2000 cases will be used for final evaluation and testing. To do this, the NN saved after the fitting step is used to predict the limiting current densities using the seven input parameters of the 2000 cases. The predicted current densities are compared with the simulated current densities from COMSOL. The mean absolute percent error (MAPE) is saved as a measurement of the effectiveness of the fitted NN.

First, networks with three hidden layers are used and the number of neurons in the first, second, third layers are randomized between [1, 20], [1, 50], and [1, 20], respectively. To do this, a subroutine is coded to perform the NN fitting, NN prediction, and MAPE computation. A main program randomizes the sizes of three layers, calls the subroutine, and saves the MAPE with each layer configuration. Figure 6 shows the MAPE for neural networks with different numbers of neuron in the three layers. The overall prediction is much better than the GAM, with the MAPE down to between 2.9 and 3.0%. Figure 6d shows the relationship between the limiting current density predicted by the NN and the ones by COMSOL simulation for the 2000 testing cases. For this particular example, the three hidden layers in the network have 7, 16, and 3 neurons, respectively, resulting in a MAPE of 3.0% and a maximum APE of 16.2%.

It is evident from Fig. 6a that the first layer is the most critical layer. A first layer with three neurons or less results in much poorer prediction and a 3% MAPE is never achieved regardless of the number of neurons in subsequent layers. Furthermore, the quality of network prediction has a wide variation with MAPE up to 16% for most networks when the number of neurons in the first layer is less than 10. On the contrary, MAPE is almost always less than 4% when more than 10 neurons are used.

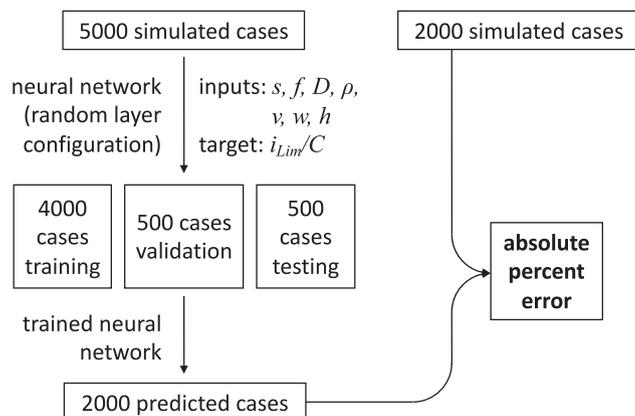
Figure 6b shows that the number of neurons in the second layer is moderately important, where at least 5 neurons are required to yield a MAPE below 3%. But the penalty for fewer neurons is less significant compared with the first hidden layer. The number of neurons in the third layer appears to be much less important and, as presented in Fig. 5c, a best MAPE of 2.9% can always be achieved regardless of the number of neuron in third layer.

Based on the understanding of neural network performance, a second type of network fitting is conducted to determine which parameter or parameters cannot be effectively described by the GAM model. From Fig. 3, it is known that the MAPE for GAM model is 12%, much higher than the 3% MAPE from NN. This study answers the question which parameters are needed in addition to the



**Figure 4.** Time evolution of velocity of vortex flow (a) in the trenches and (b) in the bulk electrolytes.

COMSOL simulation of two separate batches of randomized cases



**Figure 5.** Overall workflow of neural network fitting and evaluation. The Fitting step consists the training, validation, and testing of the network.

GAM model in order to improve the fitting and decrease the MAPE from 12% to 3%. The target of NN is still the normalized limiting current density ( $i_{Lim}/C$ ). The inputs include the limiting current density predicted by GAM model in Eq. 15 and a subset of the 7 parameters. The external Matlab code randomizes the size of three neuron layers between [1, 10], [1, 25], and [1, 10], respectively, creates a random subset of the 7 parameters, calls the neural network subroutine for fitting and evaluation, and saves the MAPE with the specific subset of parameter.

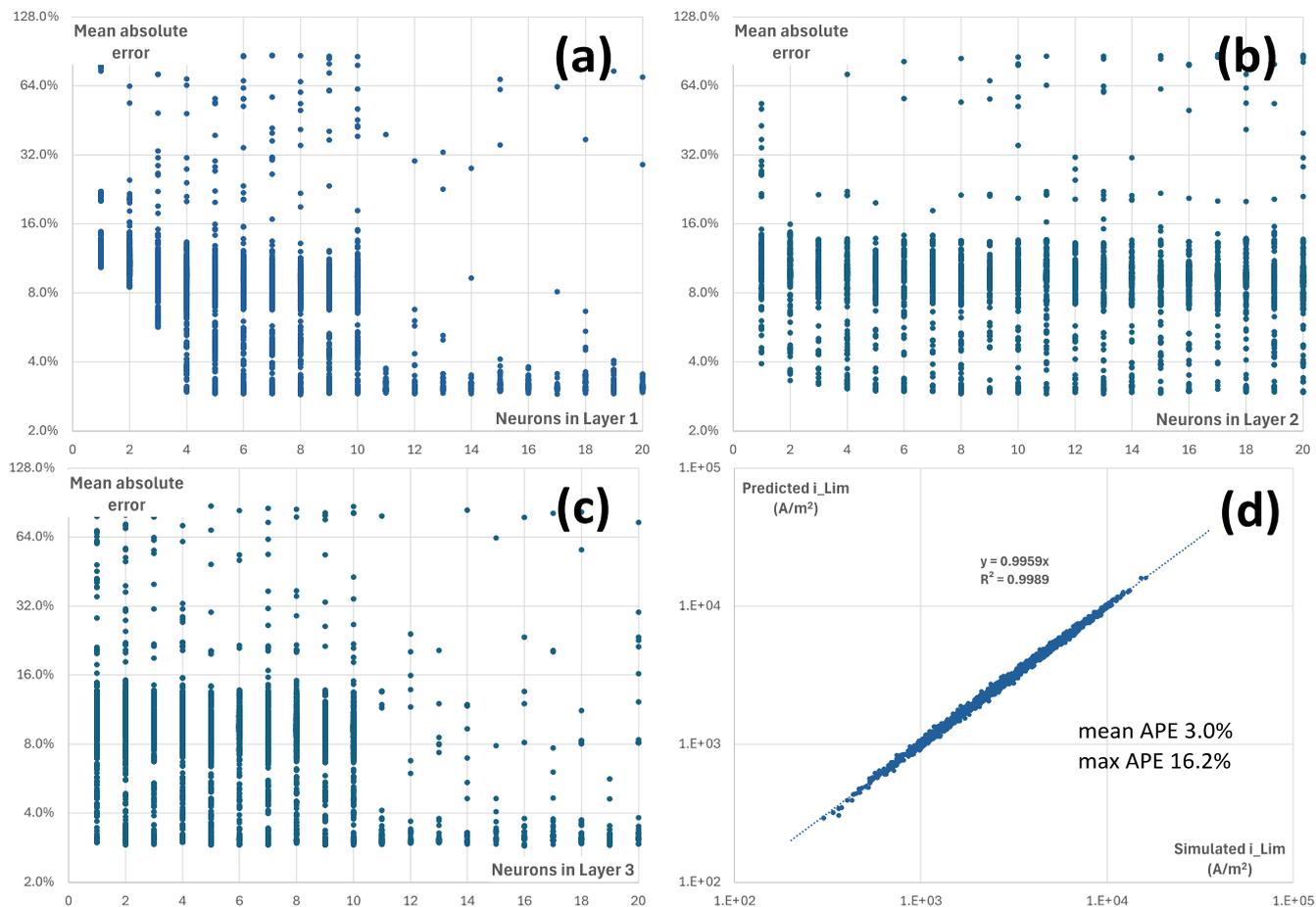
Figure 7 shows the MAPE when only 4 or 3 of the 7 parameters are used. In other words, 3 of the 7 parameters are excluded in Figs. 7a and 4 parameters are excluded in 7b. The networks cannot achieve the same accuracy of prediction (2.9%) when 5 parameters are excluded. The total possible combination for 3 or 4 out of 7 parameters are  $C_3^7 = C_4^7 = 35$ . 4000 networks with randomized neurons in the three layers are used in each study, i.e., 114 cases per combination. The vertical grid lines in Fig. 7 approximately correspond to each combination. It is evident in Fig. 7b that the only case (green) that can achieve 2.9% MAPE is when  $D$ ,  $w$ ,  $h$  are all included or when only  $s$ ,  $f$ ,  $\rho$ , and  $\nu$  are excluded. The best four cases (green) with the same 2.9% MAPE in Fig. 7a are when only a subset of the four elements  $s$ ,  $f$ ,  $\rho$ , and  $\nu$  are excluded from the inputs. On

the other hand, Fig. 7a also shows the worst prediction (red cases) results from the exclusion of any three of the four parameters,  $f$ ,  $D$ ,  $w$ , and  $h$ .

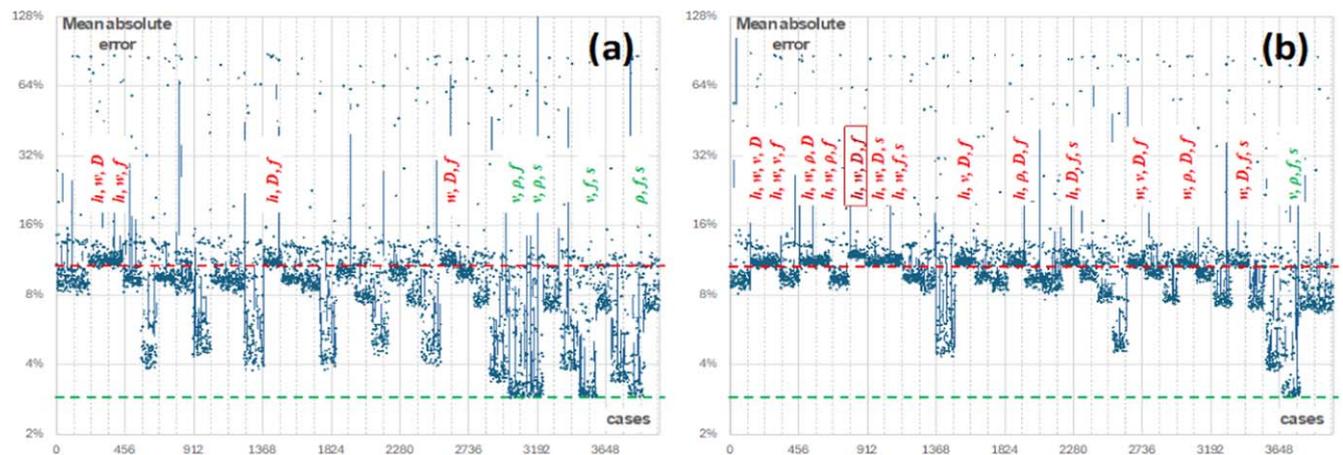
Based on these observation, the impacts of stroke length  $s$ , density  $\rho$ , and kinematic viscosity  $\nu$  appear to be sufficiently described by the GAM model. These parameters become redundant in the presence of the GAM prediction in Eq. 15. On the other hand, the diffusivity  $D$ , trench width  $w$ , and trench depth  $h$  are essential parameters in addition to the GAM prediction. They have to be present in the input in order to improve the prediction by NN. The paddle frequency,  $f$ , appears to be in between. While the exclusion of frequency in addition to the redundant parameters does not compromise the accuracy of network prediction, the exclusion of it together with the essential parameters makes the prediction worse. This is shown as the crimson case in Fig. 7b. When a combination of redundant and essential parameters are excluded from the input, the prediction performance of the NN is intermediate between the best green cases and worst red cases.

Network fitting is also carried out to explore in what form the parameters are present in the correlation. A first quick comparison is carried out to determine the effectiveness of using the parameters directly or using them in the logarithmic form. In this study, all neural networks consist of only 1 layer. The same procedure is carried out, using the first batch of 5000 cases to fit the NN and the second batch of 2000 cases to evaluate the NN and determine the MAPE. Figure 8 shows the best MAPE as a function of the number of neurons in the single layer. It is evident that the logarithmic inputs are much more efficient in predicting the mass transport behavior, resulting in a much lower MAPE regardless of how many neurons are used. It is also worth noting that the neural network with one single neuron and logarithmic inputs is similar to the GAM model fitting, where the output is a linear combination of the 7 logarithmic inputs as shown in Eq. 15. This is indeed confirmed by the similar MAPE (around 12%) between GAM fitting and the single neuron network in Fig. 8.

A further comparison is performed using NN with 3 layers and up to 4 neurons in each layer. While more neurons are beneficial for a better prediction and a lower MAPE, simple networks with less neurons are used in this study to highlight the impacts of different input formats. In this set of studies, five inputs are used including the current density predicted by GAM model ( $i_{Lim\_GAM}$ ) in Eq. 15 and the frequency ( $f$ ), diffusivity ( $D$ ), trench width ( $w$ ), and trench depth ( $h$ ). As discussed in Fig. 7, these four additional parameters are sufficient to improve the GAM prediction achieving a best MAPE of 2.9%. The



**Figure 6.** The relationship (a)–(c) between mean absolute percent error and the number of neurons in different hidden layers in the neural network; and (d) between the limiting current density from COMSOL simulation and predicted using a neural network with 7, 16, and 3 neurons in the 1st, 2nd, and 3rd hidden layers, respectively.

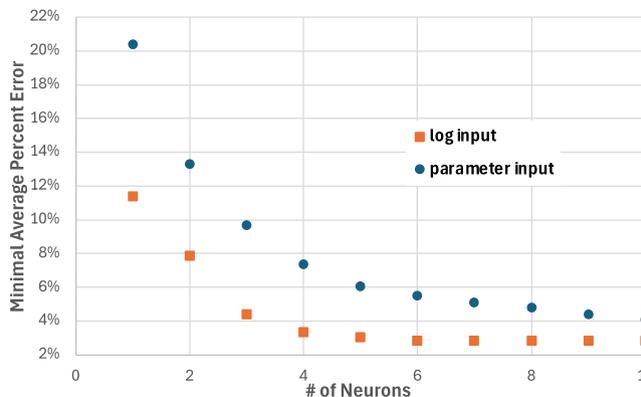


**Figure 7.** The performance of neural network prediction when the inputs are the GAM fitted limiting current density and a subset of (a) 4 or (b) 3 parameters out of 7 parameters. The notation in the figures represent the parameters that are excluded from the inputs. The colors of the notation match the two MAPE lines representing the best and worst prediction.

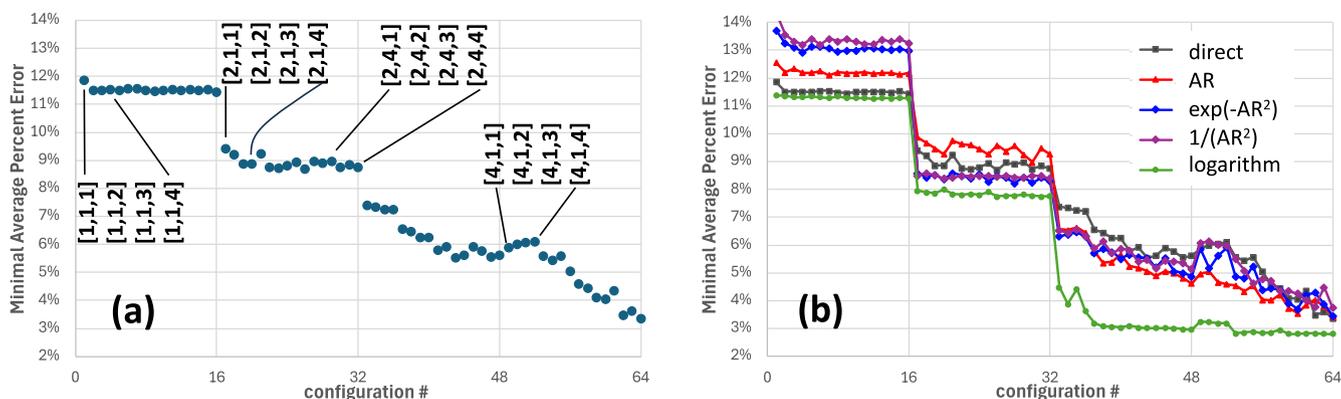
purpose of this study is to use arithmetic operators to pre-convert parameters as inputs to evaluate the efficacy of such converted parameters in improving the GAM prediction.

Figure 9 shows the best MAPE for the networks with different configurations and different input formats. The network configurations are ordered by the number of neurons in the 1st, 2nd, and then the 3rd layer. The total number of network configurations is  $4^3$  and each configuration is repeated for about 30 cases (2000 cases in

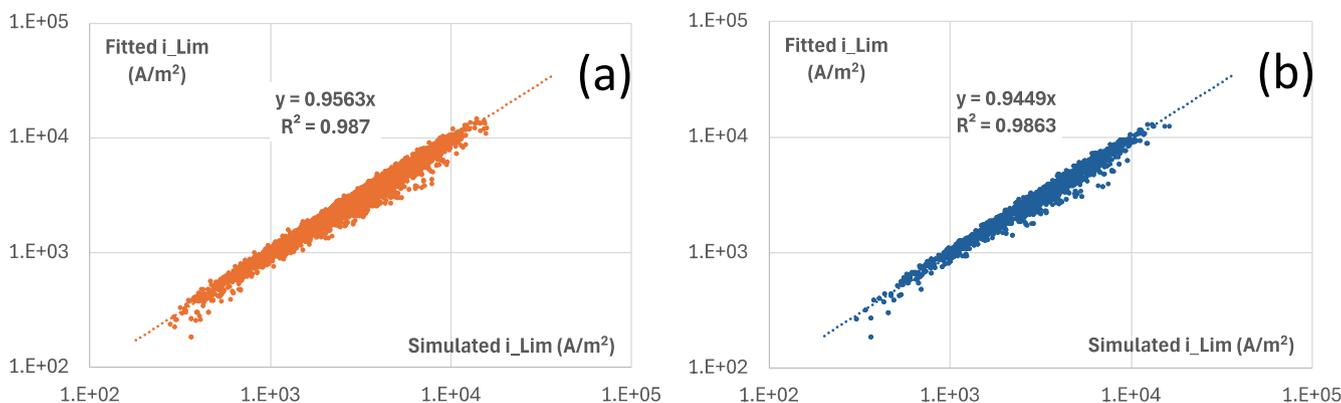
total). Figure 9a shows the best MAPE for each NN configuration when the four parameters are used directly as the inputs. The NN performance is the same as the GAM when the first layer has only one neuron. This can be expected because all the other parameter inputs are forced to fold into the single neuron in the first layer regardless of the configuration of subsequent layers. While the network performance improves when the first layer increases to two neurons, the performance is still largely independent of the second



**Figure 8.** The best MAPE using neural networks comprising only 1 hidden layer as a function of the number of neurons in that single layer. The blue points represents NNs with the parameters used directly as the inputs. The orange points represents NNs with the logarithmic parameters used as the inputs. The single neuron with logarithmic inputs has a minimal APE of 12%, similar to GAM model.



**Figure 9.** The MAPE at best fitting for neuron networks comprising 3 hidden layers and up to 4 neurons for each layer (64 configurations) with the inputs including the GAM model prediction ( $i_{Lim\_GAM}$ ). (a) Four additional parameters, paddle frequency ( $f$ ), ionic diffusivity ( $D$ ), pattern width ( $w$ ), and pattern depth ( $h$ ) are also included in inputs directly. The notation in brackets represents the number of neurons in the three hidden layers. (b) A comparison between (black) direct parameter inputs, with the depth ( $h$ ) replaced by (red) aspect ratio ( $AR = h/d$ ), by (blue) exponential decay of aspect ratio ( $e^{-AR^2}$ ), by (purple) power decay of aspect ratio ( $AR^{-2}$ ), and (green) with all inputs in logarithm forms.



**Figure 10.** The relationship between numerically simulated and fitted limiting current using the two-step transport model and GAM fitting: (a) 5000 cases for fitting and (b) 2000 cases for testing.

and third layers. However, when the first layer has three and four neurons, the performance not only further improves but also becomes more dependent on the other two layers. A lowest MAPE of 3.4% is obtained with the [4, 4, 4] network configuration.

The velocity profile and the mass transport in the trench is expected to be highly dependent on the depth of trench or the aspect ratio of the trench. Figure 9b shows the same performance comparison as Fig. 9a except that the depth is replaced by a function

of aspect ratio. Such replacements do not change the degree of freedom among the inputs while providing arithmetical operations in advance. It is evident that such replacement does not significantly change the network performance. When the first layer has only 1 neuron, such replacement compromises the fitting, suggesting the aspect ratio and the functions of aspect ratio are less effective than the depth itself as an NN input. When the first layer has three or four neurons, aspect ratio becomes slightly more effective as an input,

suggesting that the aspect ratio may be present in the correlation equations of the mass transport.

Another case in Fig. 9b is where the four additional parameters are used in their logarithmic forms. When the first layer has only one neuron (configuration 1 to 16), it does not result in a better performance than GAM. This is expected because the GAM is already a linear combination of the logarithmic forms of these parameters. Therefore, no improvement is expected from combining more logarithmic parameters with the GAM prediction into the only neuron. However, when the number of neurons in the first layer increases, the performance rapidly improves, consistent with the observation for single layer network in Fig. 8. Furthermore, the performance remains independent of the size of the second and third layers. Because linear combination of logarithmic parameters are simply products of power functions, the recursive linear combination operations in the subsequent layers appear to be redundant despite of the sigmoid normalization at each layer. While the detailed insights of neuron networks remain unknown, this comparison suggests that the difference between the mass transfer rate and the rate predicted by the GAM model,  $i_{Lim} - i_{Lim\_GAM}$ , is also a combination of power functions and can be described using another GAM model. In other words, the GAM model in Eq. 15 can be further improved by adding another GAM model.

**Corrected GAM model.**—A simplified two step mass transport physics model is proposed. The first step represents the convective transport from bulk solution to the top opening of the trench, and the second step represents the transport from the opening to the bottom electrode. The two steps are in series and the overall mass transport coefficient can be described as

$$\frac{nF \cdot C}{i_{Lim}} = \frac{1}{h_{mass}} = \frac{1}{h_{mass,1}} + \frac{1}{h_{mass,2}} \quad [22]$$

where  $h_{mass,1}$ ,  $h_{mass,2}$ , and  $h_{mass}$  represent the convective mass transport coefficients in the first step, second step, and overall process, respectively. The spherical diffusion at the trench opening is ignored. Based on the discussion in Fig. 9, both convective transport steps can be described using  $Sh = k \cdot (Re)^m \cdot (Sc)^n$ , resulting in the overall transport as a combination of two power function correlations. A multi-varient fitting is performed for this two step model using the first 5000 cases for fitting and the other 2000 cases for testing. While a true global minimum of the fitting cannot be guaranteed, this fitting has been performed using generalized reduced gradient solver with 100 starting points<sup>30</sup> and a same minimum has been reached multiple times. Figure 10 shows the relationship between numerical simulated limiting current density and the limiting current predicted with the following two correlations

$$h_{mass,1} = e^{-1.589} \cdot s^{0.399} \cdot f^{0.686} \cdot D^{0.592} \cdot \nu^{-0.232} \cdot w^{-0.101} \cdot h^{-0.052} \quad [23]$$

$$h_{mass,2} = e^{-1.979} \cdot s^{0.099} \cdot f^{0.280} \cdot D^{0.791} \cdot \nu^{-0.168} \cdot w^{1.752} \cdot h^{-2.437} \quad [24]$$

Compared with the single step GAM fitting in Eq. 15, the overall MAPE is reduced from 12.7% to 7.7% with the two step transport GAM model. It is very interesting that one of the fitted correlations, Eq. 23, strongly depends on  $s$  and  $f$ , suggesting strong influence from the paddle convection on this mass transfer step. On the other hand, this step is almost independent or very weakly dependent on the trench geometry,  $w$  and  $h$ . Therefore, this Eq. 23 describes the transport in the bulk electrolyte outside the trench. The other transport correlation in Eq. 24 describes the mass transfer rate within the trench. It is strongly dependent on the trench dimensions

and almost independent with the stroke length  $s$ . It weakly depends on the paddle frequency probably due to the time averaging across half a period as discussed in Eqs. 18 to 21. Both convection coefficient correlations are consistent with Eq. 21.

The two convective mass transfer coefficients can be further converted into the dimensionless forms as below.

$$Sh_1 \approx 0.204 \cdot (Re)^{0.686} \cdot (Sc)^{0.408} \quad [25]$$

with

$$L_{c,1} \propto s^{0.886} \cdot w^{0.312} \cdot h^{0.160} \\ Sh_2 \approx 0.138 \cdot (Re)^{0.280} \cdot (Sc)^{0.209} \quad [26]$$

with

$$L_{c,1} \propto s^{0.251} \cdot w^{-2.433} \cdot h^{3.385}$$

The dimensions of the two characteristic lengths in the two transport steps are  $[m^{1.36}]$  and  $[m^{1.20}]$ , respectively, also suggesting additional contributions from other dimensions such as the cell size and paddle geometries. It is worth noting that the characteristic length listed above for the Eqs. 14, 17, 25 and 26 may not be final because additional dimensionless number such as aspect ratio or other ratios between dimensions can be present in the mass transport correlation.<sup>21</sup> For instance, in the second mass transport step, where the trench geometry is important, the convection coefficient in Eq. 24 shows that it depends on  $w^{1.752} \cdot h^{-2.437}$ . If the aspect ratio  $h/w$  is present as a separate term in the mass transport correlation, it will change the definition of characteristic length. Furthermore, it is known from early studies<sup>20</sup> that the mass transport in paddle cells on flat surface is strongly influenced by the paddle shape, paddle size and the gap between paddle and cathode. These parameters can also pair up as additional dimensionless numbers in the correlation, further changing the characteristic length. While future studies would be required to better understand the influence of these parameters, the dimension derived above,  $[m^{1.20}]$ , remains unchanged suggesting that impacts of additional dimensions must be present in the correlation.

While a fitted neural network can predict the limiting current with a MAPE below 3%, it relies on a large number of parameters in the hidden layers without physical meaning. On the other hand, GAM model provides partial understanding of the physical process and allows analytical calculation of the limiting current. Analytical integration of the current also becomes possible to predict the metal growth in the trench. However, the accuracy of a single GAM model is much lower. The analysis of different neural networks in Figs. 8 and 9 clearly suggest that the limiting current can be more accurately predicted using a linear combination of multiple GAM models.

$$i_{Lim} = \sum_{j=1}^{n_{GAM}} GAM_j \quad [27]$$

From Fig. 8, the MAPE is expected to be about 12% for a single GAM, 7% for two GAM, and below 4% when three or more GAM are used. However, such addition of GAM can be viewed as simply allowing more fitting parameters. The analysis in this section however demonstrates a direction using NN analysis to help understand the physical process. For example, Eq. 22 provides the physical understanding of the second GAM in the model. It is well known that spherical diffusion governs the limiting current on a microelectrode, which is proportional to the ion diffusivity and inversely proportional to the electrode size.<sup>31</sup> The trench width in this study is between 20 to 150  $\mu m$ , making the spherical diffusion a nontrivial contributor to the overall mass transport and the potential mechanism for a third GAM component. The trenches are not isolated and the spherical diffusion boundary layer is periodically interrupted

by forced paddle convection in this study. This contribution is expected to be different from the classic spherical diffusion and dependent on the parameters beyond diffusivity and trench width.

### Conclusions

The mass transport behavior in patterned micro-sized trenches under reciprocating paddle agitation are systematically studied using numerical simulation and neural network analysis. A large number of numerical models are computed to cover a wide parameter space of trench geometry, electrolyte properties, and operation conditions. A correlation with power functions between the mass transport Sherwood number and dimensionless Reynolds and Schmidt numbers are obtained using generalized additive model (GAM). Neural networks are also used not only to provide much better predictions, but also to determine which parameters have additional impacts on the mass transport beyond the GAM model and in what forms these parameters take to describe these additional impacts. A two-step convective mass transport process is proposed, where the first step describes the transport outside the micro-trenches and the second step describes the transport inside the trench. An improved correlation is demonstrated for the overall mass transport rate using this two-step model.

### Acknowledgments

This work is supported by National Science Foundation through grants 1662332 and 1941820 and BASF SE. The author would also like to dedicate this work to Dr. Lubomyr Taras Romankiw, who invented the paddle cell and had been a great mentor, colleague, and friend.

### ORCID

Q. Huang  <https://orcid.org/0000-0002-1391-6531>

### References

- W. Koh, B. Lin, and J. Tai, "Copper pillar bump technology progress overview." *12th International Conference on Electronic Packaging Technology and High Density Packaging 2011* (2011).
- C. Y. Wu, C. H. Wang, K. K. Ho, K. M. Chen, P. C. Kuo, and C. L. Yang, "Chip package interaction development of flip chip CSP package with Cu pillar bump on lead for advanced node chip." *11th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT) 2016* (2016).
- K. P. S. Haesevoets, A. Radisic, and P. M. Vereecken, "Copper rich Cu1-xNix alloys ( $0.05 < x < 0.15$ ) electrodeposited from acid sulfate-based electrolyte with benzotriazole additive for microbump metallization for 3D stacked integrated circuits." *J. Electrochem. Soc.*, **166**, D315 (2019).
- L. Fu, M. Bhagavat, C. Selvanayagam, K. Leong, and I. Barber, "Cu pillar bump development for 7-nm chip package interaction (CPI) technology." *Journal of Microelectronics and Electronic Packaging*, **17**, 1 (2020).
- C. Stream, K. Emile, and T. Maggie, "Bump shape control on high speed copper pillar plating process in lead-free wafer level packaging." *4th International Microsystems, Packaging, Assembly and Circuits Technology Conference 2009* (2009).
- P. T. Lee, Y. S. Wu, C. Y. Lee, H. C. Liu, and C. E. Ho, "High-speed Cu electrodeposition and reliability of Cu pillar bumps in high-temperature storage." *J. Electrochem. Soc.*, **165**, D647 (2018).
- L. Tien-Yu Tom, W. H. Lytle, and B. Hileman, "Application of a CFD tool in designing a fountain plating cell for uniform bump plating of semiconductor wafers." *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, **19**, 131 (1996).
- M. E. Coltrin and R. J. Kee, "Unified Nusselt-and Sherwood-number correlations in axisymmetric finite-gap stagnation and rotating-disk flows." *Int. J. Heat Mass Transfer*, **102**, 122 (2016).
- A. Sonin, "Jet impingement systems for electroplating applications: mass transfer correlations." *J. Electrochem. Soc.*, **130**, 1501 (1983).
- R. Alkire and J. B. Ju, "High speed selective electroplating with impinging two-dimensional slot jet flow." *J. Electrochem. Soc.*, **134**, 294 (1987).
- J. Yang, A. Li, X. Zhu, X. Liu, and Z. Shen, "Numerical study on the hydrodynamic performance of a novel electroplating cell with jet array and crossflow." *Transactions of the IMF*, **98**, 42 (2020).
- O. N. Şara, J. Erkmen, S. Yapici, and M. Çopur, "Electrochemical mass transfer between an impinging jet and a rotating disk in a confined system." *Int. Commun. Heat Mass Transfer*, **35**, 289 (2008).
- J. V. Powers and L. T. Romankiw, *Electroplating Cell Including Means to Agitate the Electrolyte in Laminar Flow* (International Business Machine, US) (1972).
- D. E. Rice, D. Sundstrom, M. McEachern, L. Klumb, and J. B. Talbot, "Copper electrodeposition studies with a reciprocating paddle." *J. Electrochem. Soc.*, **135**, 2777 (1988).
- P. C. Andricacos, K. G. Berridge, J. O. Dukovic, M. Flotta, J. Ordenez, H. R. Poweleit, J. S. Richter, L. T. Romankiw, O. P. Schick, and F. Spera, *Vertical Paddle Plating Cell* (International Business Machine, US) (1996).
- B. Q. Wu, Z. Liu, A. Keigler, and J. Harrell, "Diffusion boundary layer studies in an industrial wafer plating cell." *J. Electrochem. Soc.*, **152**, C272 (2005).
- L. Romankiw, I. Croll, and M. Hatzakis, "Batch-fabricated thin-film magnetic recording heads." *IEEE Trans. Magn.*, **6**, 597 (1970).
- L. Romankiw and D. Thompson, "Magnetic properties of plated films." *Properties of Electrodeposits, Their Measurement and Significance* (The Electrochemical Society, Princeton, NJ) 389 (1975).
- E. E. Castellani, J. V. Powers, and L. T. Romankiw, *Nickel-iron (80: 20) Alloy Thin Film Electroplating Method and Electrochemical Treatment and Plating Apparatus*, US Patent 4102756 (1978).
- D. T. Schwartz, B. G. Higgins, P. Stroeve, and D. Borowski, "Mass-transfer studies in a plating cell with a reciprocating paddle." *J. Electrochem. Soc.*, **134**, 1639 (1987).
- G. J. Wilson and P. R. McHugh, "Unsteady numerical simulation of the mass transfer within a reciprocating paddle electroplating cell." *J. Electrochem. Soc.*, **152**, C356 (2005).
- G. Jarjoura, M. Muinonen, and G. Kipouros, "Physicochemical properties of nickel copper sulfate solutions." *Can. Metall. Q.*, **42**, 281 (2003).
- T. Kalliomaki, A. T. Aji, L. Rintala, J. Aromaa, and M. Lundstrom, "Models for viscosity and density of copper electrorefining electrolytes." *Physicochemical Problems of Mineral Processing*, **53**, 1023 (2017).
- Y. Umetsu, Q. Su, and K. Tozawa, "Density and viscosity coefficients of acidic copper sulfate solutions containing added impurities." *Shigen-to-Sozai*, **105**, 693 (1989).
- R. A. Noulty and D. G. Leait, "Diffusion in aqueous copper sulfate and copper sulfate-sulfuric acid solutions." *J. Solution Chem.*, **16**, 813 (1987).
- G. Graham, L. P. Chua, B. Buckalew, T. Ponnuswamy, and S. Mayer, "Within die coplanarity improvement strategies for electroplated Cu pillars." *IEEE 20th Electronics Packaging Technology Conference (EPTC). 2018. IEEE* (2018).
- T. L. Bergman, *Fundamentals of Heat and Mass Transfer* (Wiley, New York) (2011).
- J. Welty, G. L. Rorrer, and D. G. Foster, *Fundamentals of Momentum, Heat, and Mass Transfer* (Wiley, New York) (2014).
- E. L. Cussler, *Diffusion: Mass Transfer in Fluid Systems*. (Cambridge University Press, Cambridge, United Kingdom) (2009).
- D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview." *Soft Computing*, **22**, 387 (2018).
- A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications* Vol. 2(Wiley, New York) (1980).