PHYSICS-CONSTRAINED NEURAL DIFFERENTIAL EQUATIONS FOR LEARNING MULTI-IONIC TRANSPORT

Danyal Rehman
Center for Computational Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
drehman@mit.edu

John H. Lienhard
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
lienhard@mit.edu

ABSTRACT

Continuum models for ion transport through polyamide nanopores require solving partial differential equations (PDEs) through complex pore geometries. Resolving spatiotemporal features at this length and time-scale can make solving these equations computationally intractable. In addition, mechanistic models frequently require functional relationships between ion interaction parameters under nano-confinement, which are often too challenging to measure experimentally or know *a priori*. In this work, we develop the first physics-informed deep learning model to learn ion transport behaviour across polyamide nanopores. The proposed architecture leverages neural differential equations in conjunction with classical closure models as inductive biases directly encoded into the neural framework. The neural differential equations are pre-trained on simulated data from continuum models and fine-tuned on independent experimental data to learn ion rejection behaviour. Gaussian noise augmentations from experimental uncertainty estimates are also introduced into the measured data to improve model generalization. Our approach is compared to other physics-informed deep learning models and shows strong agreement with experimental measurements across all studied datasets.

1 INTRODUCTION

Highly-selective polyamide membranes are used ubiquitously across the separations industry to recover valuable metals, such as lithium and cobalt (DuChanois et al., 2023). Owing to the rapid growth of the electric vehicle industry, the demand for these metals is expected to double by 2025 and quadruple by 2030 (Sovacool et al., 2020). To meet this increasing demand, optimizing the selectivity of polyamide membranes across diverse water sources is essential and of substantial industrial interest. Models that accurately predict ion transport and selectivity without the need for expensive experiments can play a significant role in achieving these objectives (Hegde et al., 2022).

Continuum dynamics models are frequently used to describe the underlying laws of physical phenomena using partial differential equations (PDEs) (Li et al., 2020b; Raissi et al., 2019; Brunton et al., 2016). Solving these PDEs yields high solution accuracy but often at the expense of computational cost (Karniadakis et al., 2021). With ion transport across highly-selective polyamide nanopores, which are often fabricated using chaotic chemical processes like interfacial polymerization, classical numerical methods encounter highly irregular and non-homogenous boundary conditions that necessitate high-resolution discretizations to resolve (Ritt et al., 2020b; Jimenez-Solomon et al., 2016). Additionally, at these length-scales, ion interactions become non-negligible and require functional relationships between interaction parameters and the governing PDE – under nano-confinement, these typically invariant parameters start to diverge from bulk values due to the spatial orientation constraints (Epsztein et al., 2020; Abbaszadeh et al., 2023; Geise et al., 2014)[1].

To address these concerns, physics-informed neural solvers that abstract out the nature of these complex boundaries and pore geometries can prove meaningful in deriving accurate and generalizable models for ion transport under nano-confinement. In this work, we propose a physics-constrained architecture that combines neural differential equations with hard inductive

[1] Additional research pertaining to deep learning for PDEs and continuum ion transport models is covered in Appendix A.1.

biases to account for charge conservation (Chen et al., 2018). In addition, we leverage established mechanistic solvers and independent experimental data to train the neural architecture to improve generalization across diverse concentration inputs (Bowen & Welfoot, 2002). Gaussian noise augmentations are also introduced to the measured data to improve solver performance. This effort is the first attempt at using physics-constrained deep learning to learn complex ion transport dynamics and rejection behaviour across highly-selective polyamide membranes.

2 Proposed Architecture

Neural Ordinary Differential Equations The continuous dynamics of the hidden layers, $\mathbf{h}$, in the neural network are parameterized using a first-order ordinary differentiatial equation (ODE):

$$\frac{d\mathbf{h}(J_v)}{dJ_v} = f_\theta(\mathbf{h}(J_v), J_v; \theta) \tag{1}$$

for flux, $J_v = \{0 \dots \mathcal{J}_v\}$, and $\mathbf{h} \in \mathbb{R}^d$, where d denotes the maximum number of ionic species present across all datasets. Additionally, f_θ: $[0, \mathcal{J}_v] \times \mathbb{R}^d \to \mathbb{R}^d$. To account for mixtures with different ions in the training and test data, masking is applied to $\mathbf{h}(0)$. Here, outputs of the hidden layer correspond to scalar concentrations, $\mathbf{h}(J_v)$. Additionally, $\theta \in \Theta$, where Θ represents some finite dimensional parameter space (Chen et al., 2018). By learning the derivative of the hidden layer output, concentrations are uniformly Lipschitz continuous in $\mathbf{h}(J_v)$ and continuous in J_v, enabling pre-training on classical continuum models (Kidger, 2022).

In addition to masking, polynomial positional encodings are used. The embeddings are concatenated with the masked concentration vector prior to being passed into the neural network. To integrate over the neural ODE, we use the Dormand–Prince 5(4) numerical method (Calvo et al., 1990) and backpropagate through the solver using the continuous adjoint method (Chen et al., 2018).

The network is comprised of five linear layers and $\tanh(\cdot)$ non-linearities applied to each output. Following the last linear layer, no point-wise activations are used. The network is trained using Adam with a batch size of 32 and an initial learning rate of 10^{-3} (Kingma & Ba, 2014). The learning rate is halved every 200 epochs for a total of 1000 epochs. For all experiments conducted, we evaluate the hidden state dynamics and their derivatives on the GPU using PyTorch, which were obtained from Python's scipy.integrate package (Virtanen et al., 2020; Paszke et al., 2019).

2.1 Inductive Biases

Charge Conservation Electroneutrality in ion-fluid systems is a conservation law that necessitates a solution's net charge remain neutral under equilibrium conditions (Alkhadra et al., 2022). Within the nanopores, local electroneutrality can break down (de Souza et al., 2021); however, in the bulk fluid, $\forall J_v$, the following constraint holds:

$$\sum_{j=1}^{d} z_j \mathbf{h}_j(J_v) = 0 \tag{2}$$

where $z \in \mathbb{Z}^d$ is a vector of valences. To encode electroneutrality into the neural network as a hard constraint, we evaluate the orthogonal projection of the hidden layer output: $z^\top \mathbf{h}_\perp = z^\top \mathbf{h} - z^\top \mathbf{h}_\parallel$. By using $\mathbf{h}_\perp$ instead of appending the inductive bias to the loss as a soft constraint, the model enforces inter-ionic coupling between ions, substantially improving generalization performance.

2.2 Training and Augmentations

Pre-training on Continuum Models To pre-train the neural architecture, we use simulated data generated from the well-established Donnan-Steric Pore Model with Dielectric Exclusion (DSPM-DE). We use an iterative, under-relaxed numerical scheme to solve the PDE (Geraldes & Brites Alves, 2008) (model and implementation details are provided in Appendix A.2):

$$J_i = -D_i K_{i,d} \partial_x C_i + K_{i,c} C_i J_v - \frac{K_{i,d} D_i C_i z_i F}{RT} \partial_x \psi, \qquad x \in [0, \Delta x_e] \tag{3}$$

In DSPM-DE, four latent variables are most often used to parameterize the nanoporous membrane: $\mathcal{Z} = \{r_p, \Delta x_e, \zeta_p, \chi_d\}$. We apply our previously-developed approach using global optimization with simulated annealing and the Nelder-Mead local search to regress average values of the four parameters across the training data (Rehman & Lienhard, 2022). Details are provided in Appendix A.2.

During pre-training, the d-dimensional concentration vector was sampled using low-discrepancy Sobol sequences and projected to log-space to improve model predictions at lower concentrations (Joe & Kuo, 2008; Wang et al., 2021b). The MSE loss used for pre-training is:

$$\mathcal{L}_{\text{cm}}(\mathbf{h}, \mathbf{h}^{\text{cm}}) = \frac{1}{kd} \sum_{i=1}^{k} \sum_{j=1}^{d} \left[\mathbf{h}_j(J_{v,i}) - \mathbf{h}_j^{\text{cm}}(J_{v,i})\right]^2 \tag{4}$$

Fine-tuning on Measured Data Following pre-training, the network was fine-tuned using experimental data comprising 850 ion concentration measurements. To improve generalization, we use measured uncertainties across the training data to fit Gaussian statistics to individual data points. To evaluate the loss function, values of output concentration are sampled from this distribution[2]:

$$\mathcal{L}_{\text{exp}}(\mathbf{h}, \mathbf{h}^{\text{exp}}) = \frac{1}{nd} \sum_{i=1}^{n} \sum_{j=1}^{d} \left[\mathbf{h}_j(J_{v,i}) - \mathbf{h}_j^{\text{exp}}(J_{v,i})\right]^2, \quad \mathbf{h}_j^{\text{exp}}(J_{v,i}) \sim \mathcal{N}(\mu_{ij}, \sigma_{ij}^2)\ \forall i, j \tag{5}$$

All datasets analyzed the same polyamide membrane (DuPont's FilmTec™ NF270) for ion separation (Al-Zoubi et al., 2007; Micari et al., 2020; Al-Zoubi & Omar, 2009; Epsztein et al., 2018). To weight the measured data more heavily than the simulated data during training, $n \ll k$.

3 Results and Discussion

Downstream Task Prediction To evaluate the performance of our framework, we compare ion rejection predictions with those from the continuum model and measurement data in the test set (Fig. 1A). Rejection values closer to 1 signify perfect selectivity between the ion and the polyamide membrane. Ion rejection can also undertake values below 0, which signify counter-ion entrainment through the pore to satisfy the electroneutrality condition in the product flow (Gilron et al., 2001). Negative rejection is frequently observed in ion transport across polyamide nanopores (Alkhadra et al., 2022), and here, we illustrate the neural solver's ability to capture this phenomena and generalize it more accurately than the baseline DSPM-DE method.

In Fig. 1B, a sample initial condition from the test set is propagated through the neural solver and DSPM-DE. Since the baseline method requires knowledge of the hydrated ionic radii and membrane charge (a parameter with high sensitivity to local concentration and composition) – both of which can possess large uncertainties – error in these parameters propagates into ion rejection (Rehman & Lienhard, 2022). Since the neural ODE eliminates the need for these parameters by treating the polyamide like a black-box, predictive performance is substantially improved.

Other Architectures We compare our approach to other deep learning methods and mechanistic models in Table 1. We test physics-informed neural networks (PINNs)[3] (Raissi et al., 2019), ResNets (He et al., 2015), and U-Nets (Ronneberger et al., 2015). For the mechanistic models, we use DSPM-DE as the baseline method (Bowen & Welfoot, 2002). We also evaluate the solution-friction (SF) method on the test data using ion-specific friction factors regressed from the training set (Wang et al., 2021a). Since inter-species diffusion coefficients for all ions in the training and test data were unavailable in the literature, the Maxwell–Stefan framework could not be used for comparison purposes. The projection operator, noise augmentations, and training procedure were all held constant for the comparison, as were the approximate number of parameters used in the deep learning models. Hyperparameter tuning was also performed on all the learning architectures to ensure a fair comparison between rejection predictions.

[2]Hinge loss terms based on the Hofmeister series were originally included in the loss function but provided mixed results (Somrani et al., 2013). Given that ion rejection has been seen to diverge from the Hofmeister series under certain conditions, it was removed entirely from the loss function (Luo & Wan, 2011).

[3]To evaluate the PDE loss in the PINN case, DSPM-DE was used to solve the Nernst-Planck equations.

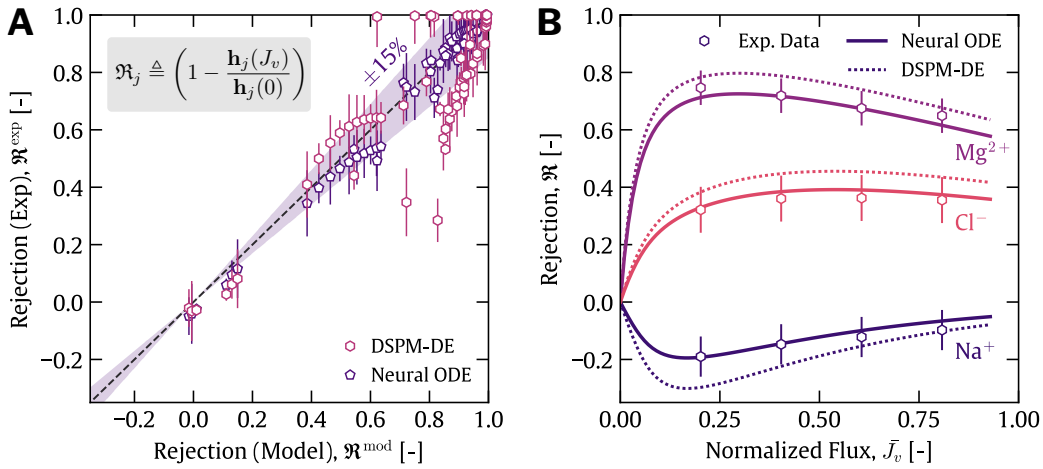

Figure 1: **Left**: Parity plot illustrating superior predictive performance of the physics-constrained neural ODE over the classical mechanistic solver across all test data. **Right**: For a given test composition, ion rejection is predicted as a function of normalized flux by the neural solver and DSPM-DE.

Table 1: Downstream task test error using alternate deep learning methods and continuum models.

	# Params	Test Error
Deep Learning Methods		
Neural ODE (Our Work)	1.05 M	7.1%
PINN (Raissi et al., 2019)	1.12 M	9.7%
ResNet (He et al., 2015)	1.23 M	9.2%
U-Net (Ronneberger et al., 2015)	1.01 M	10.2%
Continuum Models		
DSPM-DE (baseline) (Bowen & Welfoot, 2002)	4	17.8%
SF (Wang et al., 2021a)	$d+4$	24.8%

Across the test data, we find that the physics-constrained neural ODE outperforms the PINN, ResNet, and U-Net. This is largely attributed to the smooth concentrations predicted by the neural ODE, which are advantageous in bounding test error when unseen concentrations and fluxes are observed (Cuomo et al., 2022). Despite this, the other models still exhibit lower test error than the baseline DSPM-DE method, reiterating the importance of relaxing the simplifications encoded into classical mechanistic models. Similarly, in agreement with expectation, the SF method produces the largest error on the test set as a result of the overparameterized friction factors (see Appendix A.1 for details). Despite the fewer latent variables in DSPM-DE, it still exhibits higher prediction accuracy than the SF model. Overall, the neural ODE provides the best performance in predicting ion transport across the polyamide nanopores across diverse input concentrations.

4 Conclusions and Future Work

In this work, we develop the first physics-constrained neural ODE solver for ion transport across polyamide nanopores. Our solver encodes electroneutrality into the architecture and trains on a mixture of simulated data and experimental measurements augmented with Gaussian noise to achieve an average test error of 7.1%, compared to 17.8% from classical PDE solvers. The model also outperforms other deep learning models when rejection prediction is used as the downstream task. Next steps include generating new ion rejection profiles using complex mixtures to identify high-value separations using polyamide membranes. Additionally, we hope to quantify the model's predictive accuracy in the high-salinity regime for applications to metal recovery from hypersaline brines.

ACKNOWLEDGMENTS

The authors thank the Centers for Mechanical Engineering Research and Education at MIT and SUSTech (MechERE Centers at MIT and SUSTech) for partially funding this research. D.R. acknowledges financial support provided by a fellowship from the Abdul Latif Jameel World Water and Food Systems (J-WAFS) Lab and fellowship support from the Martin Family Society of Fellows.

REFERENCES

Mahsa Abbaszadeh, Madeline Garell, Ji Il Choi, Yudan Chen, Johannes Leisen, Seung Soon Jang, Yan-Yan Hu, and Marta C. Hatzell. Unraveling water and salt transport in polyamide with nuclear magnetic resonance spectroscopy. *ACS Materials Letters*, 5(0):291–298, 2023. URL https://doi.org/10.1021/acsmaterialslett.2c00932.

Habis Al-Zoubi and Waid Omar. Rejection of salt mixtures from high saline by nanofiltration membranes. *Korean Journal of Chemical Engineering*, 26(3):799–805, 2009. URL https://doi.org/10.1007/s11814-009-0133-7.

Habis Al-Zoubi, Nidal Hilal, Naif A. Darwish, and Abdul W. Mohammad. Rejection and modelling of sulphate and potassium salts by nanofiltration membranes: neural network and spiegler–kedem model. *Desalination*, 206(1):42–60, 2007. URL https://doi.org/10.1016/j.desal.2006.02.060.

Mohammad A. Alkhadra, Xiao Su, Matthew E. Suss, Huanhuan Tian, Eric N. Guyes, Amit N. Shocron, Kameron M. Conforti, J. Pedro de Souza, Nayeong Kim, Michele Tedesco, Khoiruddin Khoiruddin, I Gede Wenten, Juan G. Santiago, T. Alan Hatton, and Martin Z. Bazant. Electrochemical methods for water purification, ion separations, and energy conversion. *Chemical Reviews*, 122(16):13547–13635, 2022. URL https://doi.org/10.1021/acs.chemrev.1c00396.

Yohai Bar-Sinai, Stephan Hoyer, Jason Hickey, and Michael P. Brenner. Learning data-driven discretizations for partial differential equations. *Proceedings of the National Academy of Sciences*, 116(31):15344–15349, 2019. URL https://doi.org/10.1073/pnas.1814058116.

Richard W. Bowen and Julian S. Welfoot. Modelling the performance of membrane nanofiltration—critical assessment and model development. *Chemical Engineering Science*, 57(7):1121–1137, 2002. URL https://doi.org/10.1016/S0009-2509(01)00413-4.

Johannes Brandstetter, Max Welling, and Daniel E. Worrall. Lie point symmetry data augmentation for neural PDE solvers, 2022. URL https://arxiv.org/abs/2202.07643.

Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016. URL https://doi.org/10.1073/pnas.1517384113.

Manuel Calvo, Juan I. Montijano, and Luis Randez. A fifth-order interpolant for the Dormand and Prince Runge-Kutta method. *Journal of Computational and Applied Mathematics*, 29(1):91–100, 1990. URL https://doi.org/10.1016/0377-0427(90)90198-9.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018. URL https://proceedings.neurips.cc/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf.

Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022. URL https://doi.org/10.1007/s10915-022-01939-z.

Bram De Jaegher, Wim De Schepper, Arne Verliefde, and Ingmar Nopens. Enhancing mechanistic models with neural differential equations to predict electrodialysis fouling. *Separation and Purification Technology*, 259:118028, 2021. URL https://doi.org/10.1016/j.seppur.2020.118028.

Pedro J. de Souza, Amir Levy, and Martin Z. Bazant. Electroneutrality breakdown in nanopore arrays. *Physical Review E*, 104:044803, Oct 2021. URL https://doi.org/10.1103/PhysRevE.104.044803.

William M. Deen. Hindered transport of large molecules in liquid-filled pores. *AIChE Journal*, 33(9):1409–1425, 1987. URL https://doi.org/10.1002/aic.690330902.

Ryan M. DuChanois, Nathanial J. Cooper, Boreum Lee, Sohum K. Patel, Lauren Mazurowski, Thomas E. Graedel, and Menachem Elimelech. Prospects of metal recovery from wastewater and brine. *Nature Water*, 1(1):37–46, 2023. URL https://doi.org/10.1038/s44221-022-00006-z.

Razi Epsztein, Evyatar Shaulsky, Nadir Dizge, David M. Warsinger, and Menachem Elimelech. Role of ionic charge density in donnan exclusion of monovalent anions by nanofiltration. *Environmental Science & Technology*, 52(7):4108–4116, 2018. URL https://doi.org/10.1021/acs.est.7b06400.

Razi Epsztein, Ryan M. DuChanois, Cody L. Ritt, Aleksandr Noy, and Menachem Elimelech. Towards single-species selectivity of membranes with subnanometre pores. *Nature Nanotechnology*, 15(6):426–436, 2020. doi: 10.1038/s41565-020-0713-6. URL https://doi.org/10.1038/s41565-020-0713-6.

Geoffrey M. Geise, Donald R. Paul, and Benny D. Freeman. Fundamental water and salt transport properties of polymeric materials. *Progress in Polymer Science*, 39(1):1–42, 2014. ISSN 0079-6700. URL https://doi.org/10.1016/j.progpolymsci.2013.07.001.

Vítor Geraldes and Ana M. Brites Alves. Computer program for simulation of mass transport in nanofiltration membranes. *Journal of Membrane Science*, 321(2):172–182, 2008. URL https://doi.org/10.1016/j.memsci.2008.04.054.

Jack Gilron, Naim Gara, and Ora Kedem. Experimental analysis of negative salt rejection in nanofiltration membranes. *Journal of Membrane Science*, 185(2):223–236, 2001. URL https://doi.org/10.1016/S0376-7388(00)00639-6.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. URL https://arxiv.org/abs/1512.03385.

Varun H. Hegde, Michael F. Doherty, and Todd M. Squires. A two-phase model that unifies and extends the classical models of membrane transport. *Science*, 377(6602):186–191, 2022. URL https://doi.org/10.1126/science.abm7192.

Maria F. Jimenez-Solomon, Qilei Song, Kim E. Jelfs, Marta Munoz-Ibanez, and Andrew G. Livingston. Polymer nanofilms with enhanced microporosity by interfacial polymerization. *Nature Materials*, 15(7):760–767, 2016. URL https://doi.org/10.1038/nmat4638.

Stephen Joe and Frances Y. Kuo. Constructing Sobol sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing*, 30(5):2635–2654, 2008.

George E. Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021. URL https://doi.org/10.1038/s42254-021-00314-5.

Patrick Kidger. On neural differential equations, 2022. URL https://arxiv.org/abs/2202.02435.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014. URL https://arxiv.org/abs/1412.6980.

Dmitrii Kochkov, Jamie A. Smith, Ayya Alieva, Qing Wang, Michael P. Brenner, and Stephan Hoyer. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences*, 118 (21):e2101784118, 2021. URL https://doi.org/10.1073/pnas.2101784118.

Rajamani Krishna and Johannes A. Wesselingh. The Maxwell-Stefan approach to mass transfer. *Chemical Engineering Science*, 52(6):861–911, 1997. URL https://doi.org/10.1016/S0009-2509(96)00458-7.

Omar Labban, Chang Liu, Tzyy Haur Chong, and John H. Lienhard. Fundamentals of low-pressure nanofiltration: Membrane characterization, modeling, and understanding the multi-ionic interactions in water softening. *Journal of Membrane Science*, 521:18–32, 2017. URL https://doi.org/10.1016/j.memsci.2016.08.062.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations, 2020a. URL https://arxiv.org/abs/2010.08895.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Andrew Stuart, Kaushik Bhattacharya, and Anima Anandkumar. Multipole graph neural operator for parametric partial differential equations. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6755–6766, 2020b. URL https://proceedings.neurips.cc/paper/2020/file/4b21cf96d4cf612f239a6c322b10c8fe-Paper.pdf.

Zichao Long, Yiping Lu, Xianzhong Ma, and Bin Dong. PDE-net: Learning PDEs from data. In *International Conference on Machine Learning*, pp. 3208–3216, 2018. URL http://proceedings.mlr.press/v80/long18a/long18a.pdf.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George E. Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021. URL https://doi.org/10.1038/s42256-021-00302-5.

Jianquan Luo and Yinhua Wan. Effect of highly concentrated salt on retention of organic solutes by nanofiltration polymeric membranes. *Journal of Membrane Science*, 372(1):145–153, 2011. URL https://doi.org/10.1016/j.memsci.2011.01.066.

Gretchen M. Mavrovouniotis and Howard Brenner. Hindered sedimentation, diffusion, and dispersion coefficients for brownian spheres in circular cylindrical pores. *Journal of Colloid and Interface Science*, 124(1):269–283, 1988. URL https://doi.org/10.1016/0021-9797(88)90348-7.

Marina Micari, Dionysia Diamantidou, Sebastiaan G.J. Heijman, Massimo Moser, Amir Haidari, Henri Spanjers, and Valentin Bertsch. Experimental and theoretical characterization of commercial nanofiltration membranes for the treatment of ion exchange spent regenerant. *Journal of Membrane Science*, 606:118117, 2020. URL https://doi.org/10.1016/j.memsci.2020.118117.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf.

Kenneth S. Pitzer. Thermodynamics of electrolytes. I. Theoretical basis and general equations. *The Journal of Physical Chemistry*, 77(2):268–277, 1973. URL https://doi.org/10.1021/j100621a026.

Maziar Raissi and George E. Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357:125–141, 2018. URL https://doi.org/10.1016/j.jcp.2017.11.039.

Mazier Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. ISSN 0021-9991. URL https://doi.org/10.1016/j.jcp.2018.10.045.

Deniz Rall, Artur M. Schweidtmann, Maximilian Kruse, Elizaveta Evdochenko, Alexander Mitsos, and Matthias Wessling. Multi-scale membrane process optimization with high-fidelity ion transport models through machine learning. *Journal of Membrane Science*, 608:118208, 2020. URL https://doi.org/10.1016/j.memsci.2020.118208.

Danyal Rehman and John H. Lienhard. Global optimization for accurate and efficient parameter estimation in nanofiltration. *Journal of Membrane Science Letters*, 2(2):100034, 2022. URL https://doi.org/10.1016/j.memlet.2022.100034.

Danyal Rehman, Yvana D. Ahdab, and John H. Lienhard. Monovalent selective electrodialysis: Modelling multi-ionic transport across selective membranes. *Water Research*, 199:117171, 2021. URL https://doi.org/10.1016/j.watres.2021.117171.

Cody L. Ritt, Jay R. Werber, Mengyi Wang, Zhongyue Yang, Yumeng Zhao, Heather J. Kulik, and Menachem Elimelech. Ionization behavior of nanoporous polyamide membranes. *Proceedings of the National Academy of Sciences*, 117(48):30191–30200, 2020a. doi: 10.1073/pnas.2008421117. URL https://doi.org/10.1073/pnas.2008421117.

Cody L. Ritt, Jay R. Werber, Mengyi Wang, Zhongyue Yang, Yumeng Zhao, Heather J. Kulik, and Menachem Elimelech. Ionization behavior of nanoporous polyamide membranes. *Proceedings of the National Academy of Sciences*, 117(48):30191–30200, 2020b. URL https://doi.org/10.1073/pnas.2008421117.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation, 2015. URL https://arxiv.org/abs/1505.04597.

Markus Schmuck and Martin Z. Bazant. Homogenization of the Poisson–Nernst–Planck equations for ion transport in charged porous media. *SIAM Journal on Applied Mathematics*, 75(3):1369–1401, 2015. URL https://doi.org/10.1137/140968082.

Anissa Somrani, Ahmed H. Hamzaoui, and Maxime Pontie. Study on lithium separation from salt lake brines by nanofiltration (NF) and low pressure reverse osmosis (LPRO). *Desalination*, 317:184–192, 2013. URL https://doi.org/10.1016/j.desal.2013.03.009.

Benjamin K. Sovacool, Saleem H. Ali, Morgan Bazilian, Ben Radley, Benoit Nemery, Julia Okatz, and Dustin Mulvaney. Sustainable minerals and metals for a low-carbon future. *Science*, 367(6473):30–33, 2020. URL https://doi.org/10.1126/science.aaz6003.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17:261–272, 2020. URL https://doi.org/10.1038/s41592-019-0686-2.

Li Wang, Tianchi Cao, Jouke E. Dykstra, Slawomir Porada, Maarten P. Biesheuvel, and Menachem Elimelech. Salt and water transport in reverse osmosis membranes: Beyond the solution-diffusion model. *Environmental Science & Technology*, 55(24):16665–16675, 2021a. URL https://doi.org/10.1021/acs.est.1c05649.

Li Wang, Danyal Rehman, Peng-Fei Sun, Akshay Deshmukh, Liyuan Zhang, Qi Han, Zhe Yang, Zhongying Wang, Hee-Deung Park, John H. Lienhard, and Chuyang Y. Tang. Novel positively charged metal-coordinated nanofiltration membrane for lithium recovery. *ACS Applied Materials & Interfaces*, 13(14): 16906–16915, 2021b. URL https://doi.org/10.1021/acsami.1c02252.

Steffen Wiewel, Moritz Becher, and Nils Thuerey. Latent space physics: Towards learning the temporal evolution of fluid flow. *Computer Graphics Forum*, 38(2):71–82, 2019. URL https://doi.org/10.1111/cgf.13620.

Yinhao Zhu and Nicholas Zabaras. Bayesian deep convolutional encoder–decoder networks for surrogate modeling and uncertainty quantification. *Journal of Computational Physics*, 366:415–447, 2018. URL https://doi.org/10.1016/j.jcp.2018.04.018.

A Appendix

A.1 Relevant Work

Machine Learning for PDEs A substantial amount of literature is dedicated to using neural networks to learn PDEs (Raissi et al., 2019; Raissi & Karniadakis, 2018; Long et al., 2018; Bar-Sinai et al., 2019). These efforts include applications to fluid simulation (Wiewel et al., 2019), flame propagation (Kochkov et al., 2021; Brandstetter et al., 2022), and shallow wave dispersion (Brandstetter et al., 2022). Some approaches investigate finite-dimensional solution operators to learn the governing PDE, but intrinsically depend on domain geometry and spatiotemporal discretization (Zhu & Zabaras, 2018). Other work has considered learning infinite-dimensional neural operators for mesh-independent applications (Li et al., 2020a; Lu et al., 2021). Specific to ion transport, prior research has focused on the development of deep learning models for membrane fouling prediction (De Jaegher et al., 2021) and multi-scale transport (Rall et al., 2020); however, studies investigating multi-ion transport across polyamide nanopores remain elusive, and the primary focus of this work.

Continuum Ion Transport Models The Donnan-Steric Pore Model with Dielectric Exclusion (DSPM-DE) is one of the most frequently used models for quantifying ion transport across polyamide nanopores (Bowen & Welfoot, 2002). DSPM-DE solves the extended Nernst-Planck PDEs in conjunction with electroneutrality constraints that serve as a closure model for the electric potential. Other variants close the PDE using solutions to the Poisson-Boltzmann equations (Schmuck & Bazant, 2015). In DSPM-DE, the boundary conditions are highly simplified, assuming cylindrical nanopores and equilibrium partitioning relationships to quantify selectivity mechanisms (Rehman & Lienhard, 2022). The solution-friction (SF) model similarly solves the extended Nernst-Planck PDEs; however, it makes no assumptions about pore geometry and regresses friction factors from data to quantify restricted transport within the nanopores (Wang et al., 2021a). Although the model eliminates assumptions around pore structure, regressing friction factors often overparameterizes the model making generalization challenging. Other models like the Maxwell-Stefan framework attempt to capture inter-species coupling through experimentally-measured binary diffusion coefficients (Krishna & Wesselingh, 1997). Although measurable in bulk solutions, determining these coefficients inside the polymer matrix is often not feasible, meaning that the model often resorts to simplifications similar to those used in DSPM-DE and SF models (Rehman et al., 2021).

A.2 Mechanistic Model and Parameter Estimation

DSPM-DE solves the extended Nernst-Planck PDE inside the active layer of the polyamide:

$$J_i = -D_i K_{i,d} \partial_x C_i + K_{i,c} C_i J_v - \frac{K_{i,d} D_i C_i z_i F}{RT} \partial_x \psi, \qquad x \in [0, \Delta x_e] \tag{6}$$

Here, C_i is the solute concentration, J_i is the solute flux, and ψ is the electric potential. F, R, and T are Faraday's constant, the universal gas constant, and absolute temperature, respectively. $K_{i,c}$ and $K_{i,d}$ are convective and diffusive hindrance coefficients used to capture reductions in the bulk diffusion coefficient, D_i, under nano-confinement. Polynomial expressions for $K_{i,c}$ and $K_{i,d}$ can be derived from perturbation theory solutions to the Navier-Stokes equations assuming ions behave like hard spheres in cylindrical pores, where $K_{i,c} \in [0, 1]$ and $K_{i,d} \in [0, 1]$ (Deen, 1987; Mavrovouniotis & Brenner, 1988).

By removing hindered transport assumptions (i.e. using bulk diffusion coefficients and setting $K_{i,c} = 1$ and $K_{i,d} = 1$), the extended Nernst-Planck PDE can also be used to solve for ion transport in the boundary layer. Mass transfer correlations are also needed to calculate the Sherwood number as a function of system geometry and flow characteristics (Labban et al., 2017). The PDEs are solved simultaneously, with electroneutrality conditions applied to each spatial domain:

$$\sum_{i=1}^{d} z_i C_i = 0 \qquad \sum_{i=1}^{d} z_i C_i + \chi_d(C_i) = 0 \tag{7}$$

The first equation prescribes electroneutrality in the boundary layer film and product flow, while the second equation enforces electroneutrality inside the polyamide. $\chi_d(C)$ is the volumetric

charge density of the polyamide membrane and a function of local composition (Ritt et al., 2020a). Adsorption isotherms have been frequently used to quantify these functional dependencies.

Lastly, boundary conditions are prescribed at the solution-membrane interface. Three equilibrium partition coefficients, ϕ_{S}, ϕ_{Di}, and ϕ_{Do}, are used to evaluate ion selectivity as a function of steric, dielectric, and Donnan exclusion terms (a schematic diagram is presented in Fig. 2):

$$\frac{\gamma_i(0^-)C_i(0^-)}{\gamma_i(0^+)C_i(0^+)} = \phi_{i,\mathrm{S}}\phi_{i,\mathrm{Di}}\phi_{i,\mathrm{Do}} \tag{8}$$

Here, γ is the ion activity coefficient, which is used to capture the departure from ideality of ions in solution. In this work, the Pitzer-Kim model is used to evaluate these coefficients (Pitzer, 1973). 0^- and 0^+ denote the solution-side and membrane-side at the interface, respectively. Calculating each partition coefficient's contributions is summarized in previous work (Rehman & Lienhard, 2022).

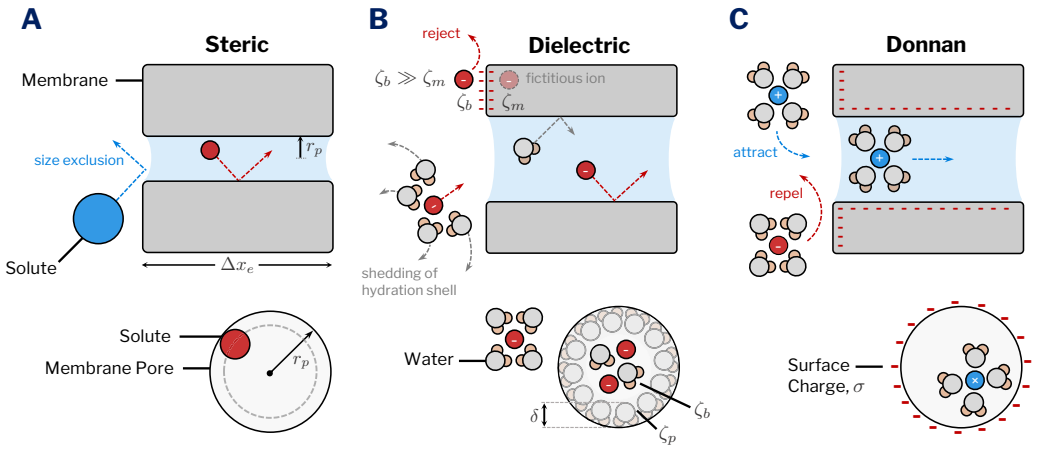

Figure 2: Physical representation of individual ion selectivity mechanisms used to quantify boundary conditions in the DSPM-DE model. **(A)** Steric rejection separates ions based on their size relative to the pore radius, r_p. **(B)** Dielectric exclusion captures the partial or complete shedding of ion hydration shells prior to entering the membrane pores with fictitious image forces created at the membrane interface. Image charges are repulsive for both cations and anions, given that the dielectric constant of the solvent, ζ_b, is larger than the dielectric constant of the membrane matrix, ζ_m. The cross section shows a thin film of water molecules with a constrained orientation aggregated near the pore walls, where the dielectric constant of the solvent is reduced to ζ_p. The thickness of the layer of water molecules is denoted by δ. ζ_b and ζ_m were set to 78.54 and 4.5, respectively, in the reported work. **(C)** The Donnan exclusion mechanism fractionates ions based on charge, e.g. negatively-charged ions are repelled by a negatively-charged membrane, whereas positively-charged ions are attracted into the pores.

To solve the above system of equations, two under-relaxation update schemes are used (Geraldes & Brites Alves, 2008). The first applies to the electric potential:

$$\psi^{(n+1)} = \psi^{(n-1)} + \eta_\psi \left[\psi^{(n)} - \psi^{(n-1)}\right] \tag{9}$$

where the superscript denotes the iteration step. The under-relaxation factor, $\eta_\psi \in [0, 1]$. Since the governing PDE can be very stiff (due to order of magnitude differences in input concentrations), convergence of the DSPM-DE method is highly sensitive to the under-relaxation parameter. A relatively low value of $\eta_\psi = 0.10$ was used across all simulations to guarantee convergence.

A second under-relaxation parameter is also required to ensure that ion concentrations con-

verge. The update step is:

$$C_i^{(n+1)} = C_i^{(n-1)} \left[1 + \eta_C \min \left(1, \left| \frac{C_i^{(n-1)}}{C_i^{(n)} - C_i^{(n-1)}} \right| \right) \right] \left(\frac{C_i^{(n)} - C_i^{(n-1)}}{C_i^{(n-1)}} \right) \tag{10}$$

where $\eta_C \in [0, 1]$. Similarly, a low value of $\eta_C = 0.175$ was set across simulations to ensure convergence. This update formulation also guarantees that $C_i^{(n+1)}$ remains non-negative at each iteration.

This procedure, in conjunction with the four latent membrane variables: $\mathcal{Z} = \{r_p, \Delta x_e, \zeta_p, \chi_d\}$, can be used to solve for output concentrations in the product flow. To quantify the latent variables, $\mathcal{Z}$, we minimize the following objective function using a hybrid global-local optimization method: simulated annealing combined with the Nelder-Mead local search (Rehman & Lienhard, 2022):

$$\arg\min_{\mathcal{Z}} \sum_{i=1}^{N_s} \sum_{j=1}^{N_w} \frac{\left[\mathfrak{R}_{i,j}^{\text{mod}}(\mathcal{Z}) - \mathfrak{R}_{i,j}^{\text{exp}} \right]^2}{\sigma_{i,j}^2} \tag{11}$$

where N_s and N_w are the total number of ions and flux measurements taken, respectively. Here, $\sigma_{i,j}^2$ is the variance estimate across experimental trials. Using this approach, in Table 2, we summarize the converged latent parameters using all data points in the training set. These values were used to generate simulated data needed for pre-training the physics-constrained neural ODE.

Table 2: DSPM-DE parameters used for pre-training the physics-constrained neural solver.

r_p [nm]	Δx_e [µm]	ζ_p [-]	χ_d [mol/m^3]
0.51	1.27	43.56	−51.23