Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2022

Machine Learning-Driven Surrogate Models for Electrolytes

Tong Gao
Michigan Technological University, tongg@mtu.edu

Copyright 2022 Tong Gao

Recommended Citation

Gao, Tong, "Machine Learning-Driven Surrogate Models for Electrolytes", Open Access Dissertation, Michigan Technological University, 2022.
https://doi.org/10.37099/mtu.dc.etdr/1472

Follow this and additional works at: https://digitalcommons.mtu.edu/etdr
Part of the Statistical, Nonlinear, and Soft Matter Physics Commons

MACHINE LEARNING-DRIVEN SURROGATE MODELS FOR ELECTROLYTES

By

Tong Gao

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Applied Physics

MICHIGAN TECHNOLOGICAL UNIVERSITY

2022

© 2022 Tong Gao

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Applied Physics.

Department of Physics

Dissertation Advisor: *Dr. Issei Nakamura*

Committee Member: *Dr. Ranjit Pati*

Committee Member: *Dr. Will Cantrell*

Committee Member: *Dr. Susanta Ghosh*

Department Chair: *Dr. Ravindra Pandey*

Contents

List of Figures . . . . **xi**

List of Tables . . . . **xxi**

Preface . . . . **xxv**

Acknowledgments . . . . **xxvii**

Abstract . . . . **xxix**

1 Introduction **1**

1.1 Lithium-ion batteries 1

1.1.1 Lithium-ion batteries development 1

1.1.2 Lithium dendrite in lithium-ion batteries 2

1.1.3 Tip effect and electric shielding 3

1.2 Soft-matter materials and related machine learning research 4

1.2.1 Introduction to soft materials 4

1.2.2 Related machine learning studies 7

1.2.3 Significance of machine learning on soft-matter physics . . . 9

1.3 Thesis Outline 10

2 Methodology 13

2.1 Diffusion-limited aggregation model 13

2.1.1 Lithium dendrite growth 15

2.1.2 Ion movement 19

2.2 Supervised learning 20

2.2.1 Artificial neural network 20

2.2.1.1 Early Stopping 23

2.2.1.2 Loss function 23

2.2.1.3 Activation function 25

2.2.1.4 Batch normalization 27

2.2.2 Ensemble neural network 28

2.2.2.1 Linear regression combination 29

2.2.2.2 Bootstrap aggregating (bagging) 30

2.3 Unsupervised learning 31

2.3.1 K-means clustering 31

2.3.2 Agglomerative clustering 32

3 Lithium Dendrite Growth in Highly Concentrated Ions 35

3.1 Introduction 36

3.2 Computational methods 40

3.3 Results and discussion 42

3.3.1 Mechanism of dendrite suppression 42
3.3.2 Symmetric ions with $d_+ = d_-$ 45
3.3.2.1 Brief summary 51
3.3.3 Asymmetric ions with $d_+ = 2d_-$ 52
3.3.4 Asymmetric ions with $2d_+ = d_-$ 55
3.4 Experimental Observation 57
3.5 Conclusion 58

4 Surrogate Model for Prediction of Lithium Dendrite Growth . . 61

4.1 Introduction 61
4.2 Computational methods 62
4.2.1 Data 62
4.2.2 Design of ENN 63
4.3 Results and discussion 66
4.4 Conclusion 70

5 Ensemble Neural Network for Predicting Dielectric Constants of Pure Solvents 71

5.1 Introduction 71
5.2 Computational methods 75
5.2.1 Data 75
5.2.2 Design of sub-NN 76
5.3 Results and discussion 79

5.3.1 Reorganizing ENN with 3 subNNs 82

5.3.2 Effect of combination method 84

5.4 Conclusion 85

6 Ensemble Neural Network for Predicting Dielectric Constants of Salt-doped Solvents 87

6.1 Introduction 87

6.2 Computational methods 88

6.2.1 Data 88

6.2.2 Design of sub-NN 89

6.3 Results and discussion 92

6.3.1 Effect of batch normalization on prediction 94

6.3.2 Effect of bootstrap aggregating (bagging) on prediction . . . 96

6.3.3 Effect of combination method 98

6.4 Conclusion 99

7 Unsupervised Learning on Cellulose Solubility in Ionic Liquids . 103

7.1 Introduction 103

7.2 Computational methods 104

7.2.1 Data 105

7.2.2 Design of ENN 106

7.3 Results and discussions 108

7.4 Conclusion 111

8 Summary and Future Perspectives **113**

8.1 Summary 113

8.2 Future perspectives 116

References **119**

A List of Selected Publications **131**

B Letters of Permission **133**

List of Figures

1.1 Lithium metal electrodeposition during the charging process in a lithium-ion battery. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 2

1.2 Charge tends to accumulate at great curvature regions. The electric field strength is larger in magnitude at locations of greater curvature. 3

2.1 2D schematic illustration of the lattice model. The blue circles describe the lithium dendrites, whereas the white circles describe dendrite candidates (or solvents) and salt ions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 15

2.2 Lattice occupancy of asymmetric particles in the lattice model. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 17

2.3 General structure of artificial neural network with two hidden layers. 21

2.4 Typical training and validation losses vs. epochs. 24

2.5 Different types of activation functions. 25

2.6 Basic ensemble neural network diagram. 29

2.7 Diagram of bootstrap aggregating (bagging). 30

2.8 K-means clustering algorithm. 32

2.9 Agglomerative clustering algorithm. 33

3.1 Ion distributions from the electrodes for the dendrite-free simulation. The diameter of the cation (anion) is 2 Å. (a) Distance from the anode vs. the number N_+ (N_-) of the cations (anions). $N = 11$ and the lattice occupancy of the ions is 76 %. (b) Distance from the anode vs. the ratio of N_+ to N_-. $N = 51$ and the lattice occupancy of the ions is 60 %. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 41

3.2 Snapshots of dendrite growth process in a simulation animation. From left to right, dendrite grows as the MC step increases until a short-circuit. 42

3.3 Schematic illustration of lithium deposition: (a) tip effect and (b) electrostatic shielding due to electric-field screening. Note that the actual electrostatic fields do not align parallel but are in highly anisotropic directions due to the dendrites. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 43

3.4 Snapshot of the potential map for 1 % salts. The color bars indicate the potential difference from the anode potential. The brighter the color, the larger the potential. The dielectric constants of the electrolytes are (a) $\varepsilon_r = 80$ and (b) $\varepsilon_r = 20$. Blue and yellowish-green particles represent the cations and anions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 44

3.5 Effects of the ionic charge, the dielectric constant on dendrite growth for the symmetric ions ($d_+ = d_- = 2$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 45

3.6 Effects of the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($d_+ = d_- = 2$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 48

3.7 Snapshots of the dendrite structure with different η values for 60% symmetric ILs and the dielectric constant $\varepsilon_r = 20$. The light blue and blue points correspond to the cations and anions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 49

3.8 Effects of the ionic charge and the parameter η on dendrite growth for the symmetric ions ($d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 50

3.9 Effects of the ionic charge, the dielectric constant on dendrite growth for the asymmetric ions ($d_+ : d_- = 2 : 1$). The x-axis is the lattice occupancy of the salts or uncharged particles. $\langle H \rangle$ designates the average height calculated after 2×10^7 MC steps. The applied voltage is 0.5 V. The dielectric constants are $\varepsilon_r= 80$. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 53

3.10 Effects of the parameter η and the applied voltage on dendrite growth for the asymmetric ions ($d_+ : d_- = 2 : 1$). The x-axis is the lattice occupancy of the salts or uncharged particles. $\langle H \rangle$ designates the average height calculated after 2×10^7 MC steps. The applied voltage is 0.5 V in (a)-(b). The dielectric constants are $\varepsilon_r= 20$. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. . . . 54

3.11 Effects of the ionic charge, the dielectric constant, the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($2d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 55

3.12 Effects of the ionic charge, the dielectric constant, the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($2d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 56

3.13 Lithium dendrite growth. (a) Original Li foil, (b) lithium deposits under 1M $LiPF_6$ in EC/DMC, and (c) lithium deposits under 1M $LiPF_6$ in EC/DMC + 10% [BMIM][TFSI]. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 57

4.1 Architecture of subNN for lithium dendrite growth prediction. . . . 63

4.2 Typical types of activation functions. 64

4.3 Comparison between the MC simulation and No.5 sub-NN in Table 4.6 for the average height and the aspect ratio. The solid lines indicate the prediction from the sub-NN in the range of 1% to 80% volume fraction of salts. The squares indicate the statistical averages calculated from the same datasets of the MC simulations as those used in Fig. 3.3c-d and Fig. 3.4a-d. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 68

4.4 Comparison between the MC simulation and ENN prediction for the average height and the aspect ratio. The symbols and lines are defined as in Fig. 4.3. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry. 69

5.1 The sub-NN structure for the pure solvents. The inputs are the dipole moment μ, the solvent diameter σ, the number density ρ, and the LJ parameter ϵ_{LJ} for the solvent-solvent interaction. The output is dielectric constant ε_r. 77

5.2 Dielectric constants predicted by No.1 ENN for different pure solvents with ϵ_{LJ}=0.01 (squares), 0.05 (circles), and 0.10 (triangles). The solid lines indicate the prediction from the ENN, solid symbols indicate the statistical averages calculated from the MD simulations. Solvents are: (a)water, (b)MeOH, (c)EtOH, (d)acetone, (e) 1-Propanole, (f)DMSO, and (g)DMF. Reproduced from Ref. [2] with permission from the Materials Research Society. 81

5.3 Dielectric constants predicted by No.3 ENN consisting of three poorly trained sub-NNs (sub-NN1, sub-NN2, and sub-NN4). The solid lines indicate the prediction from the ENN, solid symbols indicate the statistical averages calculated from the MD simulations. Solvents are: (a)water, (b)MeOH, (c)EtOH, (d)acetone, (e)1-Propanole, (f)DMSO, and (g)DMF. 83

5.4 Dielectric constants of salt-free solvents predicted by ENN1 consisting of four sub-NNs combined through a neural network. Reproduced from Ref. [2] with permission from the Materials Research Society. 84

6.1 The sub-NN structure for NaCl-containing water. The inputs are the salt concentration c_s and the LJ parameter ϵ_{LJ} for the ion-ion interaction. 90

6.2 Comparison between the ground truth and the ENN prediction for the dielectric constant. Only BN was used. The solid lines indicate the predictions from the ENN for different salt concentrations c_s and different LJ parameters between the ions. "LJ" designates $\epsilon_{LJ} ion, ion$, where $\epsilon_{LJ}^{(ion,ion)}=\epsilon_{LJ}^{(+,+)}=\epsilon_{LJ}^{(+,-)}=\epsilon_{LJ}^{(-,-)}=\epsilon_{LJ}^{(water,+)}=\epsilon_{LJ}^{(water,-)}$. The symbols indicate the ground truth (the average over 2100 configurations at each data point). Reproduced from Ref. [2] with permission from the Materials Research Society. 94

6.3 Efficacy of BN with (a) $\epsilon_{LJ} = 0.01$ kcal/mol, (b) $\epsilon_{LJ} = 0.10$ kcal/mol, and (c) $\epsilon_{LJ} = 0.50$ kcal/mol. The solid and dashed lines indicate the ENN predictions with BN and without BN, respectively. The grey symbols indicate the average over the 40 configurations at each data point, whereas the standard deviations are represented by error bars. The ENNs with BN improves on training. Reproduced from Ref. [2] with permission from the Materials Research Society. 97

6.4 BN + bagging (solid line) and bagging only (dashed line) with (a) $\epsilon_{LJ} = 0.01$ kcal/mol, (b) $\epsilon_{LJ} = 0.10$ kcal/mol, and (c) $\epsilon_{LJ} = 0.50$ kcal/mol. The symbols indicate the ground truth. The performance of the bagging-only method is tantamount to or somewhat exceeds that of BN-only method. Reproduced from Ref. [2] with permission from the Materials Research Society. 99

6.5 Dielectric constants of salt-doped solvents predicted by ENN1 consisting of sub-NNs combined through a neural network. Reproduced from Ref. [2] with permission from the Materials Research Society. 100

7.1 Ternary diagram of cellulose solubility based on experimental data. The color bar indicates cellulose solubility (wt%). 105

7.2 DNN structure for predicting cellulose solubility. This DNN has five input neurons and one output neuron. The DNN1, DNN2 have the identical structure with three input neurons ([Cl], [MeP] and [OAc]). 106

7.3 ENN structure for predicting cellulose solubility.This NN has five input neurons and one output neuron. The DNN1, DNN2 have the identical structure with ENN but only three input neurons ([Cl], [MeP] and [OAc]). 107

7.4 Ternary diagram of cellulose solubility based on ENN predictions. The color bar indicates cellulose solubility (wt%). 109

7.5 The boundary for solubility predicted by K-means clustering, where (a) K = 2 and (b) K = 5. 110

7.6 (a) A dendrogram for agglomerative clustering. The red line is where to cluster, the number of interceptions indicates the number of clusters. In the diagram, 2 interception represents 2 clusters. (b) The boundary for solubility predicted by agglomerative clustering, which has 2 clusters. 110

7.7 (a) A dendrogram for agglomerative clustering. The red line is where to cluster, the number of interceptions indicates the number of clusters. In the diagram, 5 interception represents 5 clusters. (b) The boundary for solubility predicted by agglomerative clustering, which has 5 clusters. 111

List of Tables

1.1 Experimental values of the dielectric constants ε_r for common solvents. 6

4.1 Input variables used in this study 63

4.2 Activation function structures of ENNs. 65

4.3 Activation function structures of subNNs in No.4 ENN (bold one). . 65

4.4 Hyper-parameters and their values or functions in the ENN model. 66

4.5 Training, validation, and test errors of different ENNs. The MAPE is shown in parentheses (Lower is better.) 66

4.6 Training, validation, and test errors of the sub-NNs in No.4 ENN (bold one in Table 4.5). The MAPE is shown in parentheses. 67

5.1 Examples for input variables and output for pure solvents. 76

5.2 Hyper-parameters and their values or functions in the ENN model. 78

5.3 Training and validation errors (MSEs) of sub-NNs for pure solvents. The MAPE is shown in parentheses. The topology of the sub-NNs is 4-30-30-30-30-30-30-30-30-30-30-30-1, and the sequence of the activation functions consists of X-Y-X-Z-X-Y-X-Z-X-Y-X-X, where X=ReLU, Y=sine, and Z=Gauss. 79

5.4 Training and validation errors (MSEs) of different ENNs for salt-free solvents. The MAPE is shown in parentheses. 79

6.1 Examples for input variables and output for salt-doped solvents. . . 88

6.2 Hyper-parameters and their values or functions in the ENN model. 91

6.3 ENN training errors and validation errors. The topology of sub-NNs are all 2-50-50-50-50-1. The MAPE is shown in parentheses. 92

6.4 Sub-NNs training error and validation error of sub-NNs of No.1 ENN (bold one in Table 6.3). The topology of sub-NNs are all 2-50-50-50-50-1. The MAPE is shown in parentheses. 92

6.5 MSEs (MAPEs in parentheses) of the prediction of the ENNs for NaCl-containing water with and without BN, relative to the average of the dielectric constant calculated from the poor training data sets and the ground truth. These results were calculated from Fig. 6.2 and Fig. 6.3. 95

6.6 MSEs (MAPEs) of the prediction of the ENNs for NaCl-containing water with and without BN under bagging, relative to the dielectric constant calculated from the poor training data sets and the ground truth. These results were calculated from Fig. 6.4. 98

7.1 Hyper-parameters and their values or functions in the ENN model. 108

Preface

This dissertation is an original intellectual product of the author, Tong Gao. Three topics have been studied with the machine learning-driven surrogate models including Li-ion batteries, pure solvents, salt solutions, cellulose and ionic liquids mixtures in soft matter.

Chapter 3 and 4 are reproduced from the published paper (*Mol. Syst. Des. Eng.*, 2022, 7, 260-272) with the permission of the Royal Society of Chemistry. I conducted the modeling, simulations, data collection, and analysis under the supervision of Dr. Issei Nakamura. Dr. Issei Nakamura was involved in the analysis of the results and contributed to the manuscript edits. Dr. Reza Shahbazian-Yassar and his research group were involved in experimental results that supporting computational results.

Chapter 5 and 6 are reproduced from the published paper (*MRS Communications*, 2022, 12, 966–974) with the permission of the Materials Research Society. I conducted the modeling, data collection, and analysis under the supervision of Dr. Issei Nakamura. Dr. Issei Nakamura, Dr. Amalie Frischknecht and Dr. Mark Stevens were involved in the analysis of the results and contributed to the manuscript edits.

Acknowledgments

Words cannot express my gratitude to my supervisor Dr. Issei Nakamura for his invaluable patience and guidance. I also could not have undertaken this journey without my commitee members, Dr. Ranjit Pati, Dr. Will Cantrell, and Dr. Susanta Ghosh. Additionally, this endeavor would not have been possible without the physics department chair, Dr. Ravindra Pandey, and all generous support from the Physics department. I am also grateful to my group members for their help, feedback, and moral support. Thanks should also go to Dr. Gowtham, my teaching assistant instructor Wil Slough, the physics graduate program assistant Megan Johnson and the office manager Claire Wiitanen.

Thanks High-Performance Computing Shared Facility (SUPERIOR) in Michigan Technological University, Dr. Reza Shahbazian-Yassar's research group from University of Illinois Chicago, Dr. Amalie Frischknecht and Dr. Mark Stevens from Sandia National Laboratories. Thanks all the funding resources including Henes Center Fellowship, Doctoral Finishing Fellowship, Research Excellence Fund of Michigan Technological, and NSF CAREER Award.

Last but not least, thanks my parents for passing me their diligent and intelligent genes, and thanks for their support and encouragement, so that I can reach this stage.

Abstract

We have developed a lattice Monte Carlo (MC) simulation based on the diffusion-limited aggregation model that accounts for the effect of the physical properties of ionic liquids (ILs) on lithium dendrite growth. Our simulations show that the size asymmetry between the cation and anion, the dielectric constant, and the volume fraction of ILs are critical factors to significantly suppress the dendrite growth, primarily due to substantial changes in electric-field screening. Specifically, the volume fraction of ILs has the optimal value for dendrite suppression. The present simulation method indicates potential challenges for the model extension to macroscopic systems. Therefore, we also develop ensemble neural networks (ENNs) in machine learning methods with training datasets derived from the MC simulations by considering the input descriptors with the dielectric constant, the model parameter for the fractal dimension of the dendrite, the volume fraction of ILs, and the applied voltage. Our ENNs can predict the highly nonmonotonic trend of the simulation results from only one-tenth of simulation runs, thus significantly reducing the required computation time.

To further examine the efficacy of our new ENN methods in practical applications, we apply ENNs to the study of the dielectric constants of salt-free and salt-doped solvents. Seven common solvents and NaCl solutions with various salt concentrations are considered examples. Despite the significant 50-time reduction in the number of training data, the predictions of the ENNs with batch normalization or bootstrap aggregating are largely consistent with the ground truths, tracing the optimal values out of statistically noisy data. Furthermore, we investigate the phase behaviors of cellulose and ILs mixtures by combining ENNs with unsupervised learning. As a result, K-means clustering and hierarchical clustering can automatically classify solubility phases and determine the boundaries of phases. Our work proves that machine learning could be a promising tool for studying soft matter systems.

Chapter 1

Introduction

1.1 Lithium-ion batteries

1.1.1 Lithium-ion batteries development

In recent decades, higher-capacity lithium-ion and lithium-metal secondary batteries have been developed. In the 1970s, British chemist M. Stanley Whittingham, an Exxon Mobil researcher at that time, proposed the first lithium battery (a forerunner to contemporary lithium-ion batteries). In 1991, the first commercial lithium-ion battery was released by Sony and Asahi Kasei. With the development of lithium-ion batteries, in 2019, three pioneers of lithium-ion batteries shared the Nobel Prize in

Chemistry. Despite four decades of study on lithium-ion batteries, the cell applications still face challenges. Thus, real-world lithium-ion battery applications serve as the inspiration for this thesis. Safety is unquestionably a high priority in the development of electric vehicles, laptops, and smartphones.

1.1.2 Lithium dendrite in lithium-ion batteries

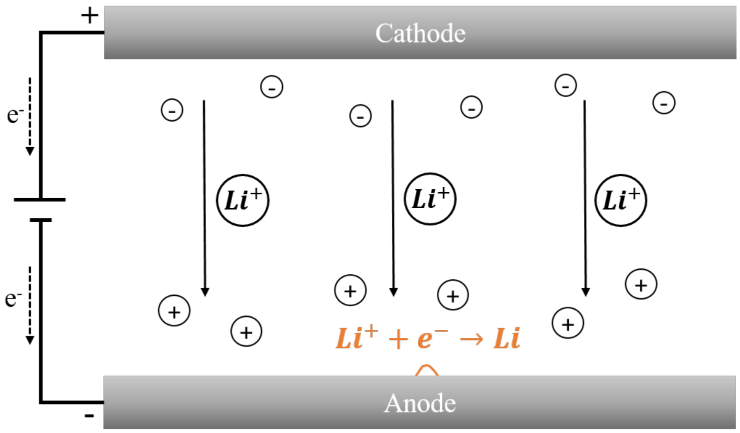


Figure 1.1: Lithium metal electrodeposition during the charging process in a lithium-ion battery. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

Lithium metal deposition is a phenomenon that occurs during the charging and discharging of lithium-ion batteries. As the battery is charged, lithium ions are released from the positive cathode, travel through electrolytes, and are intercalated into the graphene layers of the graphite (negative anode). During this charge phase, deposits solidified on the surface of the anode can form branching tree-like structures, the so-called lithium dendrite (Fig.1.1). The growth trend of the deposits is spatially uneven. Typically, needle-shaped or dendritic lithium grows on the surface of the anode over many discharge-recharge cycles [3, 4]. If dendrites grow too long, some dendrites may fall off and no longer participate in the reaction, which brings irreversible capacity loss to the battery [5, 6, 7]. Additionally, the grown dendrite may pierce the separator between the cathode and anode, forming a short circuit and potentially causing a catastrophic fire [7, 8, 9]. To maintain the safety of lithium-ion batteries, it is essential to identify the critical factors that substantially affect the process of metal solidification in electrolytes.

1.1.3 Tip effect and electric shielding

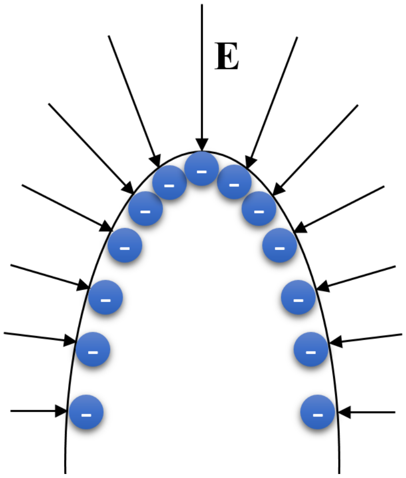


Figure 1.2: Charge tends to accumulate at great curvature regions. The electric field strength is larger in magnitude at locations of greater curvature.

The electric fields are strongest at regions along the surface where the object is most curved, this is called “tip effect”. A flat location has no curvature and has relatively weak electric fields, while a blunt point has great curvature and thus has relatively strong electric fields. Considering the diagram in Fig. 1.2, the surface that is sharply curved has strong electric fields. Likewise, the dendrite tip with a sharp edge on the metal electrode exhibits a stronger electrical field [10]. Electrostatic shielding is a phenomenon that is seen when a layer of cations is used to block the effects of an strong electric field on the curved surface. Tip effect and electrostatic shielding will be utilized to explain mechanism of lithium dendrite suppression in section 3.3.1 in detail.

1.2 Soft-matter materials and related machine learning research

1.2.1 Introduction to soft materials

Computer simulations of soft materials, including electrolytes like solvents and solutions, polymers and ionic liquids (ILs) mixtures has been an essential way for the exploration of these systems. In this thesis, more machine learning modeling of these soft-matter systems will be discussed.

Salt-free solvents and NaCl solutions. Usually, solvents are liquid substances that can dissolve a solute and become a solution. Polar or nonpolar is frequently used to categorize solvents. Polar solvents, which separate positive and negative charges at different locations in their structures, can dissolve polar compounds. Water is known to have a high dielectric constant ($\varepsilon_r = 78.36$), and water is an excellent polar solvent that is often regarded as the "universal solvent". Apart from water, other polar solvents include methanol, ethanol, acetone, 1-propanol, dimethyl sulfoxide (DMSO), and dimethylfuran (DMF). Nonpolar solvents are typically those with a dielectric constant under 15. Solvents are widely utilized as a medium for chemical reactions and phase separations. Additionally, they are necessary for the production of a wide range of goods, including cosmetics, medications, personal care products, textiles, dyes, pesticides, and so on. Sodium chloride, or NaCl, is an ionic compound in the solid state. NaCl is highly soluble in water, and water that contains NaCl is commonly called salt water in daily life.

In this work, the dielectric constants of solvents and NaCl solutions are studies as Stockmayer Fluids in Molecular Dynamics simulations and their dielectric constants are further predicted using ensemble neural networks. The dielectric constant quantifies how well a solvent can weaken the electric field that surrounds a charged particle. Since dielectric constants are connected to many significant physical and biological applications, the dielectric constant is a crucial physicochemical characteristic. Table 1.1 shows experimental dielectric constants ε_r of solvents [11], and they will be

	Water	MeOH	EtOH	Acetone	1-Propanol	DMSO	DMF
ε_r	78.36	30	24.55	21	20.1	47.2	38

Table 1.1

Experimental values of the dielectric constants ε_r for common solvents.

discussed in Chapter 5 and 6.

Cellulose. Cellulose is the most prevalent natural polymer in our surroundings, and it is a both environmentally and biologically compatible material. Cellulose is the principal structural component of plant cell walls, and it makes up to 47% of the weight of wood. Its highly organized structure gives it the desired mechanical capabilities it does, but it also makes it difficult to discover solvents that will dissolve it. Some hydrophilic ILs, like [BMIM][Cl], can dissolve cellulose. Environmental-friendly solvents of cellulose is also important for advancing the application of biopolymers.

Ionic Liquids (ILs). ILs are a class of salts that stay as liquids at low temperatures (usually less than 100 °C). They own chemical and thermal stability, nonflammability, and extremely low vapor pressure. ILs with a melting point below room temperature are called room-temperature ILs. Their enormous potential in research was just recently explored after being first found by Walden in 1914. There are 10^{12} possible ion combinations for ILs. Large, unsymmetrical ions and delocalized charges are main reasons to cause low melting points of ILs. Because there are so many different ion combinations, it is feasible to create IL solvents that have particular qualities for a given use.

ILs have a variety of physical characteristics that appeal to a wide range of applications. These desired characteristics are important since the solvent evaporation is constrained by their low vapor pressure, which eliminates the requirement for respiratory protection and exhaust systems. This characteristic also improves their recycling, which is one of the reasons they are frequently referred to as “green solvents”. As solvents, ILs appear to be highly polar due to their ionic nature, resulting in their enhanced biopolymer dissolving capacity that will be discussed in this thesis.

1.2.2 Related machine learning studies

More and more scholars use machine learning methods for research on physical and thermodynamic properties of soft materials. Chan et al. created machine learning models, which reduce computations by two orders of magnitude than atomistic models, to accurately forecast both the temperature with maximum water density and the melting point of ice [12]. Liu et al. used support vector regression for predicting the thermodynamic properties of pure fluids and their mixtures [13]. The machine learning model developed by Deng et al. can reliably forecast the dielectric constants for organic liquids and their mixtures with water [14]. To predict the free energies of solvation across a range of solvents, Hutchinson et al. implemented a machine learning framework, which outperforms the widely used traditional topology-based algorithm in terms of accuracy and the contribution of each feature to dielectric constants [15].

A machine-learning-based molecular dynamics simulation accelerates the calculation of excess chemical potentials for the solute ions Na^+ and Cl^- in the molten NaCl liquid by 3 orders of magnitude[16]. Graph neural networks are used to simulate the implicit solvent potential in molecular dynamics simulations by learning from explicit solvents behaviors in simulations [17]. Wei et al. exploited convolutional neural networks to successfully distinguish solutions containing amino acids from those containing salts in both the liquid and solid stages [18]. Alshehri et al. also used convolutional neural networks to differentiate between different salt concentrations in water by classifying photos of saline particles with various concentrations. [19]. In Tovey's work, they reported a machine learning study looking at how self-diffusion coefficients, electrical conductivity, and ion-ion distances of molten NaCl changes in high temperatures [20]. Sun et al. developed machine learning models for forecasting the refractive index and viscosity of ILs, and used ensemble learning techniques to improve the predictive performance [21]. Dhakal et el. constructed both support vector regression and artificial neural networks to investigate the ionic conductivity of imidazolium-based ILs [22]. Chandrasekaran et al. exploited a deep neural network, which is trained with more than 4500 polymers and the accompanying solvents/nonsolvents, for solvent selection in polymers [23].

1.2.3 Significance of machine learning on soft-matter physics

The discovery and invention of novel materials are ever-growing interests in materials science research, and an in-depth understanding of the atomistic or molecular mechanisms of chemical and physical properties is often required when fabricating and controlling nanotechnological devices.

There are some limitations for convention techniques in soft-matter physics. First, for some expensive materials, the number of experimental data is limited, because it is costly to collect many data from them. As a type of special and expensive materials, ILs draw international attention because of their high thermal stability, high electrical conductivity, and low viscosity. ILs are composed of pure ions and are molten salts at room temperature due to weak ion interactions. ILs are promising materials as cell electrolytes due to these properties mentioned above and are environmental-friendly solvents. Second, researchers can perform computer simulations to decrease the expense. However, a soft-matter system cannot be simply captured by a single parameter, which causes a significant cost of computations. Moreover, calculating statistical average from hundreds of, even thousands of, simulations are required.

To discover and study on novel materials in the 21st century, we should take advantage of new techniques. Machine learning is a promising tool. Machine learning implemented models have already been used in games and robots, so it is possible to be used in soft-matter physics to discover new materials or estimate the new phenomena of materials. In Chapters 4, 5, 6, and 7, machine learning methods are exploited for prediction of lithium dendrite growth in ion-containing liquids, dielectric responses of pure solvents and solutions, and phase behaviors of polymer-ILs mixtures.

1.3 Thesis Outline

Chapter 2 presents the designed diffusion limited aggregation model for lithium dendrite in lithium-ion batteries, principles of dendrite growth and ion movements. Basic definitions on artificial neural networks, ensemble neural networks, neural network training and some techniques to improve performance of neural network are also introduced.

Chapter 3 presents design of new MC model on lithium-ion batteries. This model accounts for the mechanism of dendrite growth inside the battery and simplifies the calculation. Results are discussed showing how salt ions and ILs suppress lithium dendrite growth.

Chapter 4 presents design of ENNs, evaluation of ENNs, predictions of ENNs on lithium dendrite height and aspect ratio. Results are discussed showing how an ENN outperform a single NN.

Chapter 5 presents data collection, data collection, design of sub-NNs for seven types of pure solvents with 76 training data points. In this chapter, a new architecture of sub-NN for capturing features of pure solvents' dielectric constants is proposed incorporating sandwich structure in it.

Chapter 6 presents data collection, design of sub-NNs under ENNss for salt-doped solvents. Results are discussed showing how batch normalization (BN) impacts on training ENN model with large variances training data to predict dielectric constants.

Chapter 7 presents combination of supervised learning and unsupervised learning on the cellulose solubility. Results are discussed showing how clustering algorithms efficiently determine the number of phases and predict the boundaries between phases.

Finally, Chapter 8 provides the conclusions and discusses the future work.

Chapter 2

Methodology

The current chapter presents essential knowledge of the diffusion-limited aggregation model for lattice MC simulations, NNs in supervised learning methods, and clustering techniques in unsupervised learning methods, and they are applied to solve problems in subsequent chapters.

2.1 Diffusion-limited aggregation model

Our lattice model based on the diffusion-limited aggregation (DLA) algorithm consists of a cathode at the top, an anode at the bottom, and an electrolyte between them. The gist of the lattice DLA model includes the idea that when the reaction mechanism of microscopic events is intractable, the algorithm that solidification over the coarse-grained, mesoscopic length scales occurs at the lattice points is assumed. The validity of the model is often checked against bulk values such as the fractal dimension and the scaling law of the growth rate of solidification. Here, the molar concentration of the lithium ion in Li-ion batteries typically falls in the range of 0.1 to 1 M. In this case, the volume fractions of the lithium ion fall in the range of 0.00013 to 0.0013, and the molar concentrations are significantly smaller than those of other ionic species. For example, when the volume fraction of the added salt is 0.2, the ratio of the molar concentration of the lithium ion to that of the added salt falls in the range of 0.8 to 8%. In other words, the contribution of the ionic strength of the lithium ion to those of the added salt is relatively insignificant. Thus, as an analog of the DLA model for the zinc dendrite formation [24], we treat the lithium ions dissolved in electrolytes as implicit (background) ions. The distance between each lattice point is 2 Å, and the lattice size is N×N, where N = 51. We fix the potentials of the anode (or dendrite) and cathode to 0 V and 0.5 V (unless otherwise noted), respectively. The periodic boundary condition is used in the direction parallel to the electrodes (see Fig.1.2). If we apply the periodic boundary condition, the electrode can be regarded as an infinite plate.

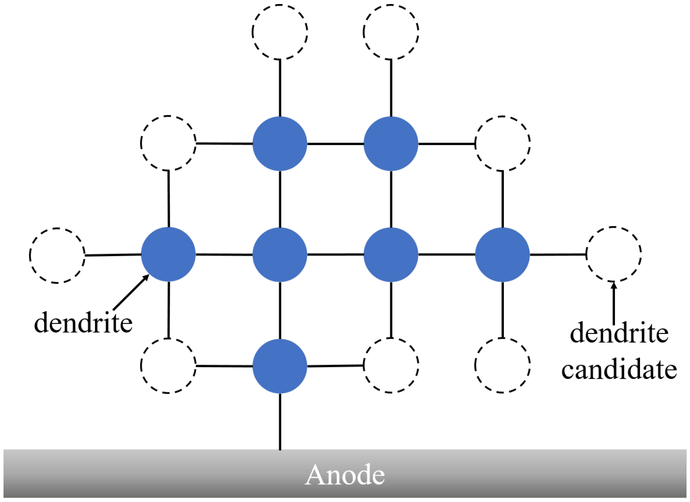


Figure 2.1: 2D schematic illustration of the lattice model. The blue circles describe the lithium dendrites, whereas the white circles describe dendrite candidates (or solvents) and salt ions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

2.1.1 Lithium dendrite growth

We illustrate our lattice system in Fig.2.1. The blue circles represent the deposited lithium metals forming the lithium dendrite. The dendrite blobs stem from the negative electrode and thus have the same potential as that of the negative electrode. The white circles indicate the empty site for the candidate for the lithium dendrite in our MC method. Salt ions can also occupy those empty sites. The electric potential $\phi_{i,j}$ at the lattice point (i,j) consists of the electric potential $\phi^E_{i,j}$ produced by the electrodes and the Coulomb potential $\phi^C_{i,j}$ produced by the ions (treated as an explicit particle) as follows:

$$\phi_{i,j} = \phi^E_{i,j} + \phi^C_{i,j} \tag{2.1}$$

Note that the growth of the lithium dendrite indicates changes in the shape of the anode surface. Accordingly, the electrostatic field between the electrodes also changes due to the change in the boundary condition of the electrostatic potential on the anode. With these altered boundary conditions due to the geometric changes, we need to solve Laplace's equation [25]

$$\Delta^2 \phi^E = 0 \tag{2.2}$$

Here, we write Eq. (2) on the lattice in the following form [26]:

$$\phi^E_{i,j} = \frac{(\phi^E_{i+1,j} + \phi^E_{i-1,j} + \phi^E_{i,j+1} + \phi^E_{i,j-1})}{4} \tag{2.3}$$

Here, $\phi^E_{i+1,j}$, $\phi^E_{i-1,j}$, $\phi^E_{i,j+1}$, and $\phi^E_{i,j-1}$ designate the electrostatic potentials of four neighboring lattice points around the (i, j) site. We note that the lattice model of Chen and Jorné included the effect of the electrical conductivity κ of water according to the modified Laplace equation, $\nabla[\kappa(\vec{r}\phi(\vec{r})] = 0$ [24]. However, the analysis of experimental data suggested that the electrical conductivity of water was about 1% of that of the dendrite, and the effect on the lattice model is insignificant. Given this situation, we do not consider the electrical conductivity κ in the present study. Concomitantly, this model simplification also enables us to explore the minimal effect to inhibit dendrite formation.

The Coulomb potential $\phi^C_{i,j}$ with the dielectric constant ε_r is given by

$$\phi^C_{i,j} = \frac{1}{4\pi\varepsilon_0\varepsilon_r}\sum_k \frac{q_k}{r_k} \tag{2.4}$$

where q_k is the k-th ionic charge and r_k is the distance between the lattice point (i, j) and the k-th ion. Fig.2.2 illustrates the lattice occupancy of the ions with diameter d = 4 Å and d = 2 Å. The ions cannot be overlapped due to the effect of the excluded volume.

A further remark about the treatment with the electrostatic interactions is discussed in the subsection, “A remark about the DLA model as a cellular automaton”.

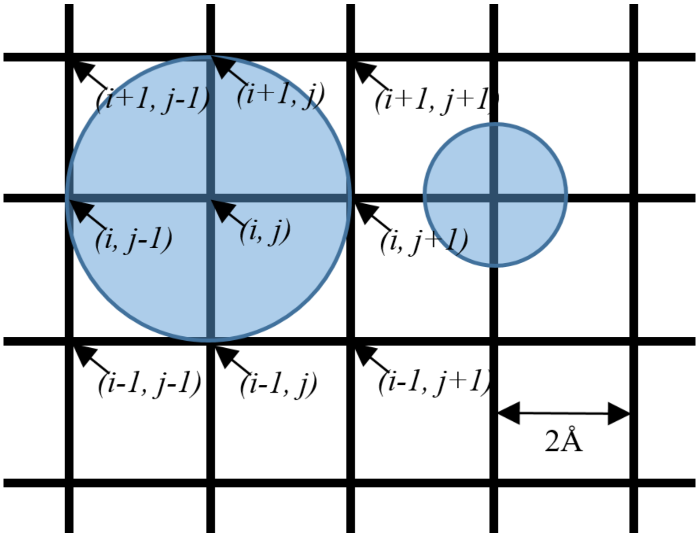


Figure 2.2: Lattice occupancy of asymmetric particles in the lattice model. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

We perform the MC update according to the local potential $\phi_{i,j}$. This algorithm includes two update processes regarding the ion configuration and the dendrite growth. Our MC simulation causes the concentration gradient and the electric field gradient along the lines of the continuum theory for ion diffusion in Ref [27]. Here, we note that the computational cost of solving Eq.(2.2) for the trial configuration for the MC update by interaction is nominal because the solution is very close to the original potential (i.e., the initial guess). The DLA algorithm of the nucleation (or the pattern formation) at the lattice point (i, j) in Refs. [24] and [26] assumes the following heuristic growth probability:

$$p_{i,j} = \frac{(\phi_{i,j})^{\eta}}{\sum_{<i,j>}(\phi_{i,j})^{\eta}} \tag{2.5}$$

Eq. 2.5 indicates the nucleation probability that the empty lattice point adjacent to the occupied lattice point is converted to the new dendrite point caused by the lithium deposition. The denominator represents the sum of the electrostatic potentials of all dendrite candidates (i.e., empty points adjacent to the occupied points). When $\phi_{i,j} \leq 0$, we set the probability to zero, assuming that the electron is unlikely to transfer from the electrode to the lithium ions in the negative potential area. The exponent η is the model parameter suggested by Niemeyer et al. [26] and associates the local potential with the nucleation probability. Experiments show that different metals have different values of the fractal dimension of the dendrites [28]. Thus, the exponent η can be determined by the fractal dimension observed in experiments.

For example, η for the zinc dendrite may be fixed to reproduce the observed fractal dimension 1.7 [24]. However, we are not aware of the fractal dimension of the lithium dendrite and therefore vary the values in the present study.

2.1.2 Ion movement

Here we write the total potential energies of the system before and after updating the positions of the ions as U_{old} and U_{new}, respectively. If the trial update decreases or does not change the potential energy, we accept the new state; otherwise, we accept the new state if a uniform random number becomes smaller than the Boltzmann distribution $e^{-\frac{(U_{new}-U_{old})}{kT}}$. The dendrite growth is significantly slower than the translational motion of the ions in the electrolytes. Accordingly, we consider the ratio of the trial update of the dendrite growth to that of the ion positions to be 1:500. It should also be noted that the present DLA algorithm is a clustering method to account for the fractal pattern of solidification and does not involve the explicit timescale of the deposition process [24, 26].

2.2 Supervised learning

2.2.1 Artificial neural network

Machine learning method is becoming more and more popular in recent decades. People expect computers to predict meaningful outputs by learning from given input data. An artificial neural network is a typical machine-learning model. An artificial neural network is composed of an input layer, several hidden layers, and an output layer. And each layer contains several neurons. The number of hidden layers and the number of neurons in each of these layers depend on the complexity of the considered system. Fig. 2.4 shows a typical ANN architecture with two hidden layers.

The number of neurons of input layer and output layer rely on the available training data. In this thesis, inputs are features affecting the system, outputs are properties we would like to look at. The details of NNs will be stated in corresponding chapters. The number of neurons in hidden layer can be one or more, and is usually larger than number of neurons in input layer and output layer. Because if the number of neurons in hidden layer is less than that in input layer, information is lost as it moves forward to the next layer.

Data is received through the input layer neurons and then transformed to the neurons

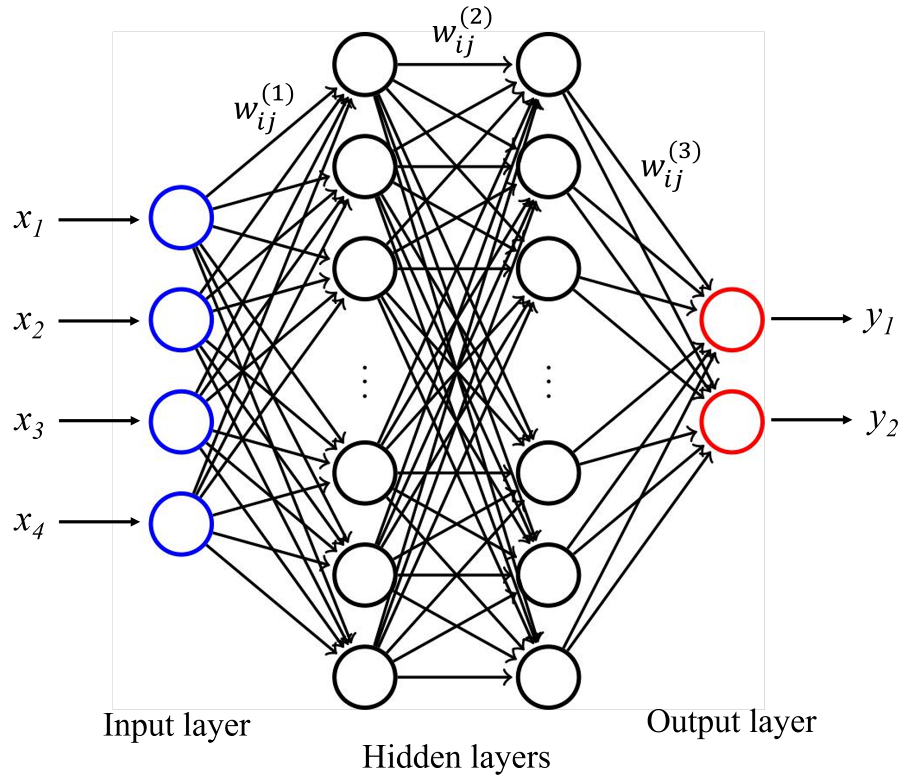


Figure 2.3: General structure of artificial neural network with two hidden layers.

in the first hidden layer through the weighted connections established between the input layer and the first hidden layer. The neurons between neighboring layers are fully connected via weights $w_{ij}^{(L)}$. The weight w_ij represents the strength of the connection between i-th neuron in a layer and j-th neuron in the next layer of the network. For example, the j-th neuron in 1st hidden layer takes input x_i from input layer, the output value of the j-th neuron, $h_j^{(1)}$, in 1st hidden layer is computed from weight $w_{ij}^{(1)}$, bias b_1, and activation function σ using Eq. 2.6. This equation explain how the incoming data (x_i) is processed by j-th node in the next layer and Fig.2.4 represents this process. Finally, the result is transferred to the neuron in the next layer.

$$h_j^{(1)} = \sigma(\sum_{i=1}^{m}(w_{ij}^{(1)} \times x_i + b_1)) \tag{2.6}$$

Neural networks with larger size (more layers and more neurons per layer) can be used to solve more complicated problems. By adjusting weights and biases using an optimizer, the artificial neural network can be used for prediction. In general, the optimizer should optimize weights and biases and then check the whole network by minimizing a loss function.

Backpropagation is the most popular optimization algorithm now, among them, gradient descent method is the most conventional method. By calculating the gradient of loss function in terms of a weight, the weight can be modified to obtain a smaller value of the loss function. Besides, stochastic gradient descent is a modified gradient descent method. Unlike gradient descent, stochastic gradient descent uses the gradient calculated from small batches of training dataset instead of the whole training dataset.

The data used to train the neural network is called training data. Training data need to be split into training dataset and validation dataset. The inputs are delivered to the neural network, and the loss function is measured by comparing the network output and truth. The training process will stop at a given number of cycles (Epochs). Because the initial weights are usually randomly arranged, the number of epochs should not be too small.

In terms of optimizing process, stochastic gradient descent is a modified gradient descent method. Unlike gradient descent, stochastic gradient descent uses the gradient calculated from small batches of training dataset instead of the whole training dataset.

2.2.1.1 Early Stopping

Early stopping is often used to prevent the network from overfitting, by stopping training once the loss function is no longer improving. We can add a patience to early stopping to make sure that the increase of the loss function is not because of fluctuation. For example, patience = 30 leads to stopping training if there is no improvement on loss function within 30 epochs.

2.2.1.2 Loss function

Loss function is a real-time function between predicted values by neural network and target values, also called error function. The goal of training a neural network is optimizing weights for connected neurons by minimizing loss function. The selection of loss function depends on the type of model, cross entropy loss function is selected while optimizing classification models, mean squared error (MSE) is most common loss function for non-classification models. The MSE of training is expressed by:

$$MSE = \frac{\sum(y-\hat{y})^2}{n} \tag{2.7}$$

where $\hat{y}$ is the output predicted by network, y is the target value, n is the number of training datasets.

The training process adjusts the unknown network parameters such as weights or biases to minimize the MSE function. As training goes, a typical loss function is like Fig. 2.4.

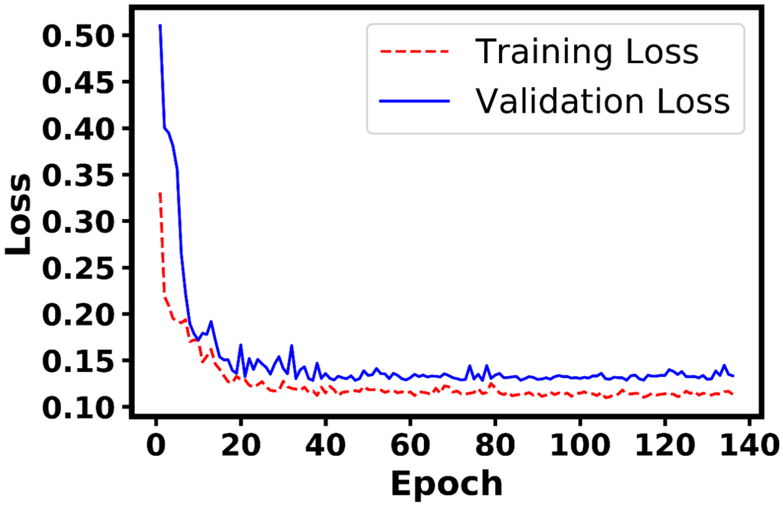
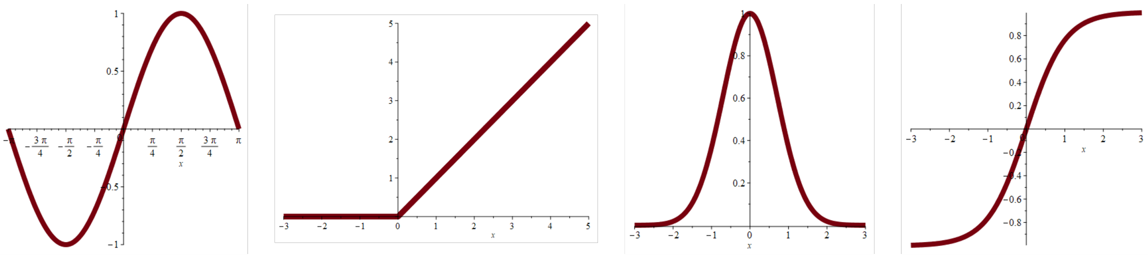


Figure 2.4: Typical training and validation losses vs. epochs.

2.2.1.3 Activation function

Popular activation functions include the piecewise function, radial basis function, the tangential function and the logistic function. In general, a non-linear function usually serve as an activation function σ. Because in some cases, a non-linear can act as a threshold, only values larger than the threshold can proceed to the next layer. Besides, non-linear functions can capture non-linear relationship between inputs and outputs, while linear activation functions can only describe linear relationship. It is obvious that NNs become more powerful with non-linear activation function. Among them,“ReLU” (rectified linear unit) function is one of the most common activation functions used in neural network. For example, a process of passing values from a layer to the next layer is applying a linear combination of weights (w), inputs and bias (b), then applying a ReLU function (Eq. 2.6), thus we get output $=$ relu($w\times$input$+b$). Other typical activation functions are listed below:



Sine function **ReLU function** **Gaussian function** **Tanh function**

Figure 2.5: Different types of activation functions.

1. Linear Activation function: This is an activation function that outputs the weighted sum of the input. This essentially acts as “no activation function”

and transfers the input directly to the output. In other words:

$$\sigma(x) = x \tag{2.8}$$

2. Logistic sigmoid: Since the derivatives of sigmoids are simple to compute, training algorithms with sigmoid functions are highly efficient. The output of this activation function ranges from 0 to 1. The logistic sigmoid is defined as

$$\sigma(x) = \frac{1}{1 + e^{-ax}} \tag{2.9}$$

3. Tangential Sigmoid: It is similar to the logistic sigmoid, but the value range of its output is from -1 to 1. The common equation for a tangential sigmoid is:

$$\sigma(x) = \frac{1 - e^{ax}}{1 + e^{ax}} \tag{2.10}$$

4. Radial Basis function: One of the most common radial basis functions is the Gaussian function, the output is the maximum, 1, when input is 0:

$$\sigma(x) = e^{-(ax)^2} \tag{2.11}$$

5. Periodic functions: One of the example is sine function.

$$\sigma(x) = sin(x) \tag{2.12}$$

2.2.1.4 Batch normalization

Here, the stochastic gradient descent, an iterative method that optimizes an objective function via the updates of the weights and biases, is often invoked for training the feedforward NNs. However, the gradient is often vanishingly small and therefore the NNs train slowly, specifically when inputs to the layers are large. To circumvent this problem, we also employ batch normalization (BN), a normalization scheme recently proposed by Ioffe and Szegedy, that accelerates the training of the NNs [29]. Since BN algorithm is new, its application to soft-matter systems is limited.

In general, normalization fixes the means and variances of the inputs to each layer. Ideally, the normalization would be performed over the complete training set. However, it is not practicable to apply normalization on the total training data especially with stochastic optimizer. As a result, the total training data are divided into batches, and normalization is limited to each batch. By dividing by the batch's standard deviation and removing the batch's mean value, BN specifically normalizes each batch of training data. Let us use b to represent a batch with m samples. The mean and variance of the batch could be written as

$$\begin{aligned}\mu_b &= \frac{1}{m}\sum_{i=1}^{m} x_i \\ \sigma_b^2 &= \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_b)^2\end{aligned} \tag{2.13}$$

where μ_b and σ_b are the mean and standard deviation of the batch, respectively. Then inputs x_i of BN layer could be normalized using

$$\hat{x}_i = \frac{x_i - \mu_b}{\sigma_b} \tag{2.14}$$

where $\hat{x}_i$ is the output of BN layer. The BN in Keras applies a transformation that maintains the mean output close to 0 and the output standard variance close to 1, mitigating the problem of internal covariate shift [30]. By applying BN, neural networks not only accelerate the training speed, but also be less likely to stuck in local minimum because it narrows the range of input data. We will explore the effect of BN on scattering data in this thesis.

2.2.2 Ensemble neural network

The ensemble neural network (ENN) consists of a set of sub-NNs with different weight initialization or other hyper-parameters in the training process to solve a problem, like wisdom of crowds.

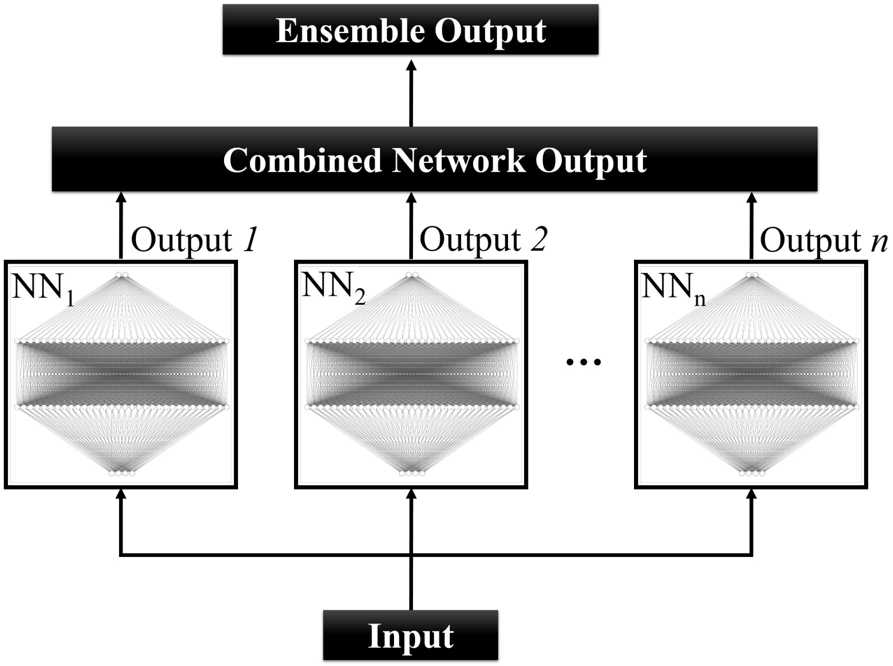

Ensemble Output
Combined Network Output
Output 1
Output 2
Output n
NN1
NN2
NNn
...
Input

Figure 2.6: Basic ensemble neural network diagram.

2.2.2.1 Linear regression combination

The outputs produced from each sub-NNs are combined using linear regression: $O_E = c_0 + c_1 O_1 + c_2 O_2 + c_3 O_3 + ...$, where $O_n (n = 1, 2, 3 \ldots)$ is the output from n-th sub-NN, O_E is the output of ENN, and $c_n (n = 0, 1, 2 \ldots)$ is the coefficient of its corresponding sub-NN. The linear regression linearly fits a linear model with coefficients c_0, c_1, ..., c_n by minimizing the sum of squared estimate of errors between the targets in the dataset and the prediction O_E from the ENN. The linear regression yields a substantially low value to $|c_n|$ when a sub-NN is poorly trained, whereas $|c_n|$ for a well-trained sub-NN becomes relatively large. Thus, the predictions of well-trained sub-NNs weigh more heavily than those of pooly trained sub-NNs. This is how sub NNs are jointly used to make a prediction. Briefly, linear regression combination fits a linear model with coefficients $c_1, c_2, \ldots, c_n$ to minimize the residual sum of squares between the targets in the dataset and the targets predicted by sub-NNs. Linear regression model will find the best fit linear line, including the optimal coefficients and intercept to minimize the error.

2.2.2.2 Bootstrap aggregating (bagging)

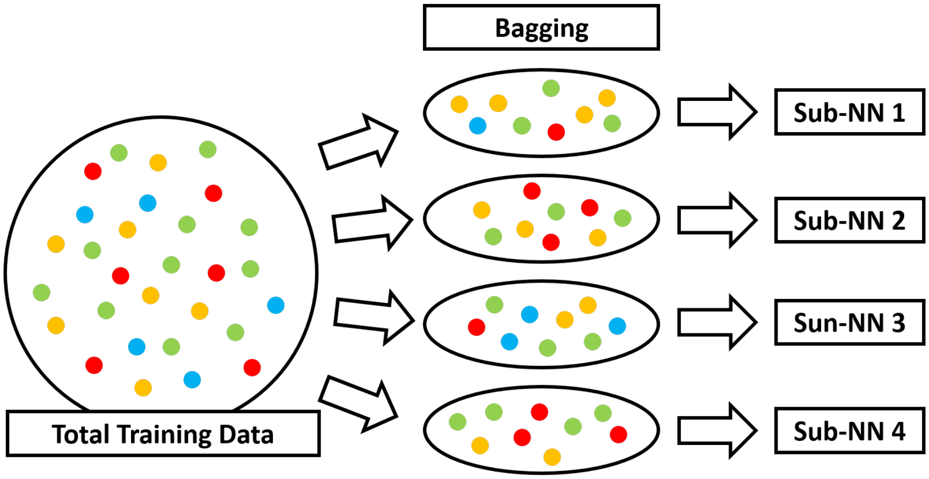


Figure 2.7: Diagram of bootstrap aggregating (bagging).

Bootstrap aggregating, also known as bagging, is a standard machine learning ensemble meta-algorithm that is commonly used to reduce variance within a noisy dataset. In bagging, samples are randomely selected from original training data and aggregate (Fig. 2.7). It is an algorithm created to enhance ensemble learning's accuracy and robustness. Meanwhile, it also reduces variance and helps to avoid overfitting.

2.3 Unsupervised learning

Unsupervised learning is discovering input-output relationships for unlabeled data. An interesting question is raised that how clusters are determined by unsupervised learning to correlate with physical and thermodynamic properties.

2.3.1 K-means clustering

The K-means clustering is one of the basic and primary used algorithms for unsupervised learning. The unlabeled dataset is divided into k separate clusters using an iterative technique, and each data point is assigned to a single cluster that shares characteristics with the others.

Data points are classified into their nearest cluster. The "K" in K-means indicates the number of clusters that be classified. The algorithm takes the unlabeled dataset as input, selects random K points as centroids, groups similar data points based on distance between them, calculates the variance and update the new centroid of each cluster, and reassigns each data point to the new closest centroid of each cluster until centroids do not change.

Fig. 2.8 illustrates how K-means clustering works. And the value of k should be predetermined in this algorithm, if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

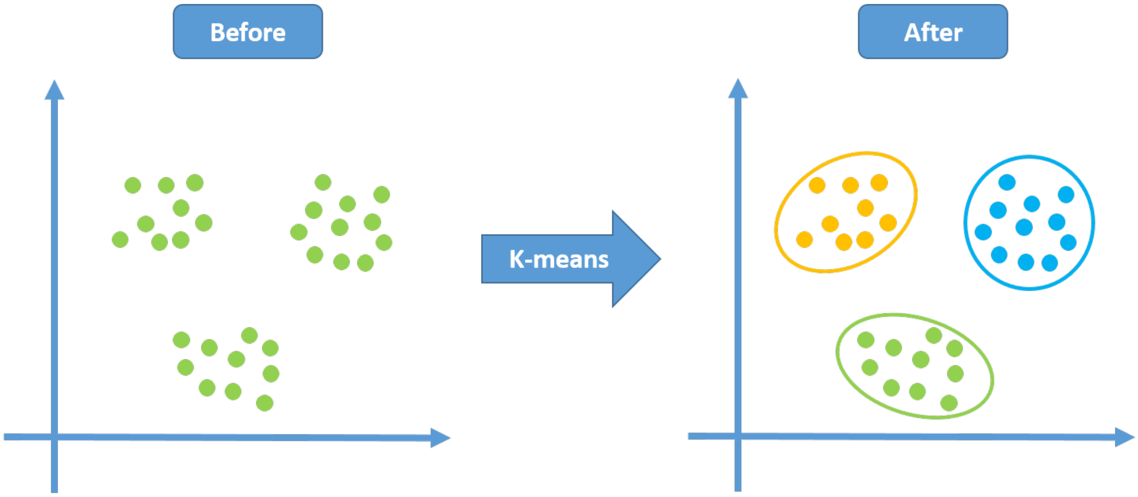


Figure 2.8: K-means clustering algorithm.

The K-means clustering algorithm start being utilized to classify soft materials based on their specific properties recently. K-means was used to separate two types of polymers, polycarbonate and polystyrene [31]. And it is also used to predict Young's modulus and ultimate tensile strength of carbon-fiber-reinforced polymers [32] and to cluster the defect and defect-free regions in glass fibre reinforced polymers with given features [33].

2.3.2 Agglomerative clustering

Agglomerative clustering is another algorithm for unsupervised learning I am going to use. Agglomerative clustering is one category of hierarchical clustering analysis that

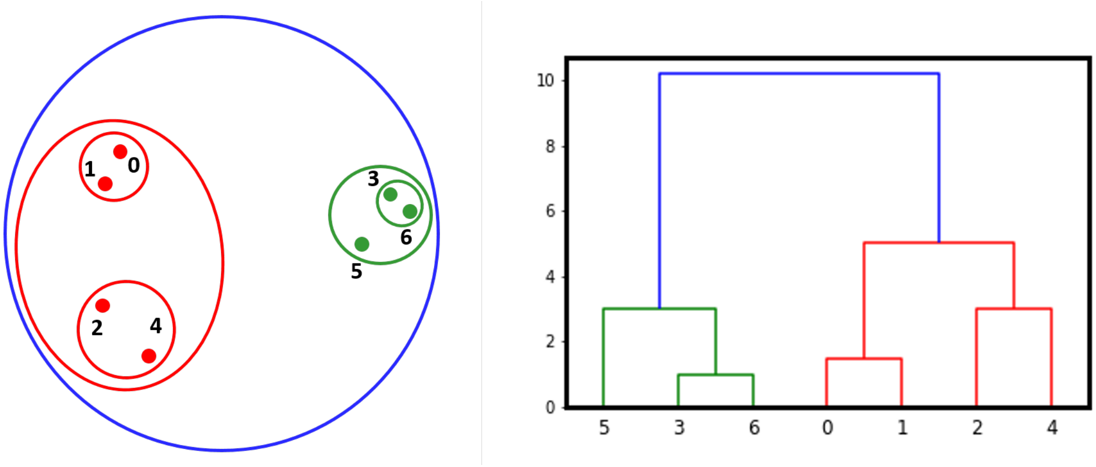


Figure 2.9: Agglomerative clustering algorithm.

seeks to build a hierarchy of clusters. It is a “bottom-up” approach: Each observation starts in its own cluster, and pairs of clusters are merged as one while going up the hierarchy.

Specifically speaking, agglomerative clustering algorithm recognizes each data point as an individual cluster at the beginning. To merge the closest pair of clusters, it calculates the distance between each of the clusters. Then, the similar clusters are iteratively merged to form a cluster according to the distance until all points are merged into one last cluster. Fig. 2.9 give an example for agglomerative clustering.

Data points in agglomerative clustering can produce a tree-like hierarchy called a dendrogram (right figure in Fig. 2.9). Numbers in x-axis represent data points in the left figure. The number of clusters will be automatically arranged when the distance of centroids is fixed, that is, we can see how many classes from a dendrogram visually.

The classification of polymer groups were investigated by Lukasiak et al. using K-means clustering and hierarchical clustering based on pattern recognition [34]. Recently, hierarchical cluster analysis was used to cluster polymers based on solid-state properties [35]. The application of unsupervised learning on soft materials research are very limited.

Chapter 3

Lithium Dendrite Growth in Highly Concentrated Ions

This chapter will exhibit effects of ion charge q, ion sizes d_- and d_+ , dielectric constant ε_r, model parameter η, and applied voltage V on lithium dendrite growth.

In the subsequent subsections, we consider two cases as follows: (1) For symmetric ions, we set the grid size (2 Å) to the cation diameter d_+ and the anion diameter d_-. (2) For asymmetric ions, we consider (d_+, d_-) = (4 Å, 2Å) and (2 Å, 4 Å).

3.1 Introduction

Specifically, a general treatment to inhibit the dendrite growth will not only help to improve the safety of Li-ion batteries but also give further insights into the electrodeposition process in the electrochemical industry [36, 37]. Nevertheless, this research objective remains challenging because (1) there exist a wide choice of system parameters, such as the size of ions, the size asymmetry between the ions in electrolytes, dielectric constant, and applied voltage, and (2) the reactions associated with the dendrite growth involve multiple length and time scales. This complexity also often complicates the computational modeling of the dendrite formation under various environments, and thus a reduction in the number of model parameters would also help in identifying the optimal design of the electrolyte systems.

Over the past few decades, many studies have aimed to improve the stability of the electrodes [38, 39, 40, 41], the qualities of electrolytes[6, 42, 43, 44, 45, 46, 47, 48, 49], and separator materials [50, 51] and the charging methods [25, 52, 53] of lithium-ion batteries. Numerous researchers have recently reported metal electrodepositions in ionic liquid (IL)-containing electrolytes [36, 43, 54, 55, 56]. Among others, an experimental study by Pearson et al. [56] suggests that the ionic strength plays an important role in causing substantial differences in dendrite growth between ILs and conventional organic liquid electrolytes. For example, the number of the nuclei increases and the dendrite growth rate decreases as the ionic strength is increased. Moreover, ILs increase the electrolyte viscosity, thus retarding dendrite formation. Similarly, our experimental observation of lithium dendrite in a combination of 1M $LiPF_6$ in EC/DMC electrolyte and 10% IL [BMIM][TFSI] also showed remarkably uniform, unconventional dendrite growth in Section 3.4. Therefore, further theoretical investigation of the effect of the ionic strength on the inhibition of dendrite formation, particularly with the visualization of dendrite growth and electrostatic potential maps, would probably provide a deeper understanding of the inhibition mechanism of dendrite growth.

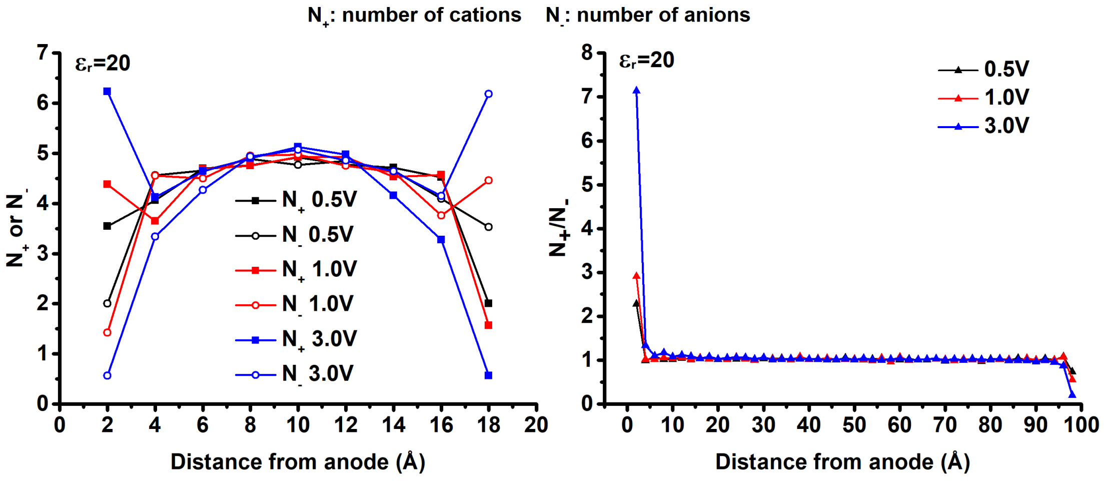
Dendrite growth involves various physical and chemical interactions that occur over multiple lengths and time scales [57]. Accordingly, various computational approaches, such as Monte Carlo (MC) simulations [25, 58, 59], molecular dynamics (MD) simulations [60], Brownian dynamics (BD) simulations [25], and phase-field models [61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76] have been developed. For example, Selis and Seminario performed classical MD simulations and investigated the relationship between charging methods and lithium dendrite formation in a nanobattery [60]. However, the time scale of the MD simulations is typically limited to the order of nanoseconds and cannot adequately account for the mechanism of mesoscopic or macroscopic metal dendrite growth associated with salt ions and ILs. BD simulations are also particle-based methods but can access longer time scales and larger length scales. Aryanfar et al. presented a novel coarse-grained MC model based on Brownian motions to explore the effect of electric current density on dendrite formation [25]. Nevertheless, BD typically assumes implicit solvents, is often inadequate to account for strong electrostatic correlations, and is normally well-defined for solid bodies with vast amounts of collisions due to surrounding molecules within the momentum relaxation time. Phase-field models are powerful tools to study the nucleation and dendrite growth that occur on relatively longer time scales and larger length scales [61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76]. Nevertheless, the model of the free energy often contains numerous adjustable parameters and may be constructed somewhat arbitrarily or empirically [66]. Thus, the development of more computationally tractable models with a relatively small number of model parameters that account for metal dendrite growth on the timescales beyond nanoseconds and the length scales beyond nanometers would also be convenient and beneficial from both the theoretical and experimental viewpoints.

Along the lines of a simulation method consisting of a minimal set of model parameters, Chen and Jorné developed a lattice simulation model for metal dendrite formation by considering zinc electrodeposition [24], given that the diffusion-limited aggregation (DLA) model captures the pattern formation of the dendrite growth. This hypothesis is legitimate because the DLA model serves as an algorithm (or a cellular automaton [77, 78] that simulates the solidification in which the microscopic reaction mechanisms are largely unknown and/or practically intractable. For example, the super-rough dynamics of tumor growth were discovered to exhibit the linear molecular beam epitaxy universality class [79, 80, 81], a feature which can be simulated by the DLA (or Eden) model [82]. In the study by Chen and Jorné, the electrodeposition and morphology of zinc dendrites in a two-dimensional battery were observed. The fractal dimension determined by digitizing photographs was 1.7 and was consistent with the value 1.71 calculated by the simulations using the DLA model originally developed for a dielectric breakdown proposed by Niemeyer et al [26]. Thus, despite the drastic simplification of the coarse-grained lattice model, the agreement between theory and experiment is remarkable. Unfortunately, however, this lattice-simulation model has not been substantially developed over the past decades, and thus the application of the simulation model to salt-doped liquids and highly concentrated ions remains limited.

In this chapter, we study the effect of small salt ions and large salt ions that mimic ILs on dendrite growth by considering the height and aspect ratio of the dendrite. We develop the DLA model on the lattice originally proposed by Niemeyer et al. to mimic lithium dendrite growth in electrolytes, mainly motivated by the model feasibility for zinc dendrite [24]. Moreover, it is well known that although the DLA model (or cellular automaton model) largely simplifies or ignores reaction mechanisms on atomistic and molecular scales, it often captures the macroscopic formation pattern of solidification. Likewise, our present cellular automaton model for electrodepositions does not contain many atomistic and molecular details regarding the reaction mechanism of the metal solidification, yet it provides a convenient tool to investigate how the electrostatic interactions affect dendrite formation with a minimal set of model parameters and thus to identify a possible mechanism for the dendrite inhibition caused by the small salt ions and ILs. Regarding the thermodynamic ensemble of ion configurations, we employ the lattice model of large ions with the excluded volume by Borukhov et al. [83]. This lattice model accounts for the electrical double layer near charged surfaces. In this chapter, we consider the dendrite growth to be limited to two-dimensional systems for computational simplicity.

3.2 Computational methods

The amount of the salts is represented by the lattice occupancy that indicates the ratio of the number of the lattice sites occupied by the salt ions to the number of the total lattice sites. The objective of the present study is to examine how the dendrite surface grows with the salt ions. Thus, in addition to the calculation of the average height of the dendrite at a given Monte Carlo (MC) step, we calculate the aspect ratio $\frac{H}{L}$ of the dendrite to characterize the dendrite morphology, where H and L designate the height and width of dendrites, respectively. For example, when the dendrite shape is relatively flat with a wider length and smaller height, the aspect ratio becomes small. We run 200 samples for each parameter space.



(a) (b)

Figure 3.1: Ion distributions from the electrodes for the dendrite-free simulation. The diameter of the cation (anion) is 2 Å. (a) Distance from the anode vs. the number N_+ (N_-) of the cations (anions). $N = 11$ and the lattice occupancy of the ions is 76 %. (b) Distance from the anode vs. the ratio of N_+ to N_-. $N = 51$ and the lattice occupancy of the ions is 60 %. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

To examine the model applicability for highly concentrated ions, we performed dendrite-free simulations with the dielectric constant $\varepsilon_r = 20$ for small [Fig. 3.1(a)] and large [Fig. 3.1(b)] systems. Overall, the results show the following two key features for large ions with excluded volume: the monolayer adsorption of the counterions near the electrodes and the recovery of the bulk density at a distance of a few monolayers from the electrodes. This fact is analogous to the results of the mean-field approximation [83] for ions with excluded volume and a MD simulation [84] for ILs in a dielectric continuum.

3.3 Results and discussion

With new MC model, we simulate the dendrite growth process.

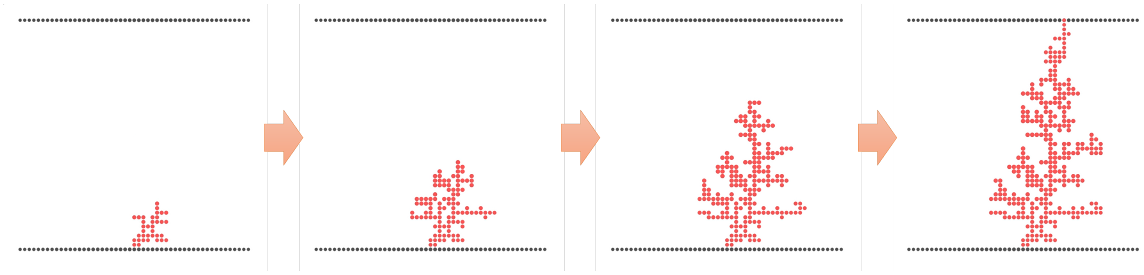


Figure 3.2: Snapshots of dendrite growth process in a simulation animation. From left to right, dendrite grows as the MC step increases until a short-circuit.

3.3.1 Mechanism of dendrite suppression

Note that a tip with a sharp edge on the metal electrode exhibits a stronger electrical field [10], which was mentioned in Section 1.1.3. Thus, lithium ions tend to be deposited preferentially around the pointy regions of the dendrite, compared to rounded or smooth regions because of "tip effect" (see Fig. 3.3a). Accordingly, the dendrite solidification near the pointy regions tends to grow rapidly and form tree-like structures.

However, when ILs are added, the cations tend to accumulate near the pointy regions

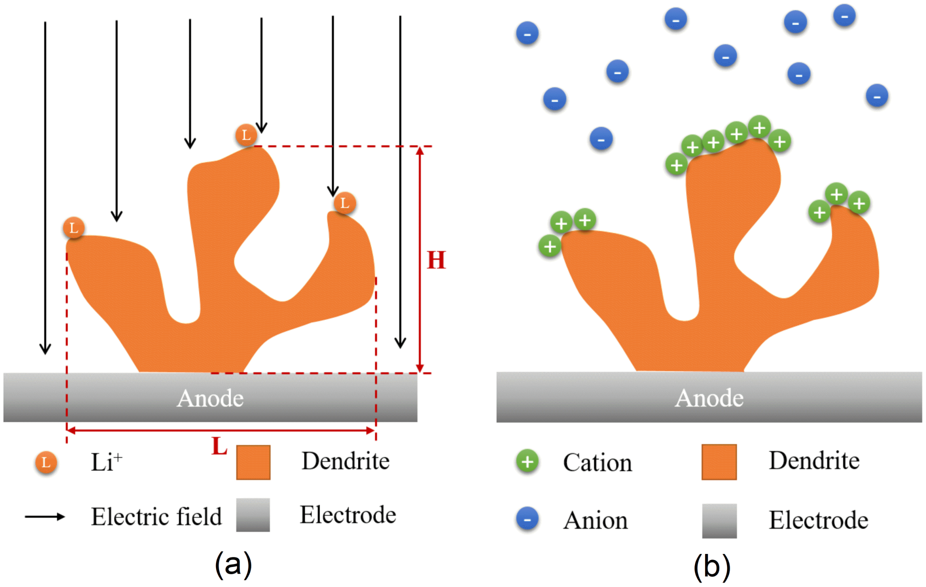


Figure 3.3: Schematic illustration of lithium deposition: (a) tip effect and (b) electrostatic shielding due to electric-field screening. Note that the actual electrostatic fields do not align parallel but are in highly anisotropic directions due to the dendrites. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

of the dendrite, forming a layer of positive charges that cause electrostatic shielding (Fig. 3.3b). These positively charged layers repel incoming Li ions from the pointy regions, and accordingly cause the Li ions to be deposited more preferentially on the adjacent area of the pointy regions. Thus, this deposition mechanism suppresses the rapid growth of the pointy regions of the lithium dendrite and leads to a relatively uniform formation of the dendrite. The anions clustered near the dendrite structure lower the local potential of the regions for dendrite candidates, and accordingly, this mechanism enhances the inhibition of the dendrite growth.

Thus, if the dielectric constant of the electrolyte changes, the distribution of cations

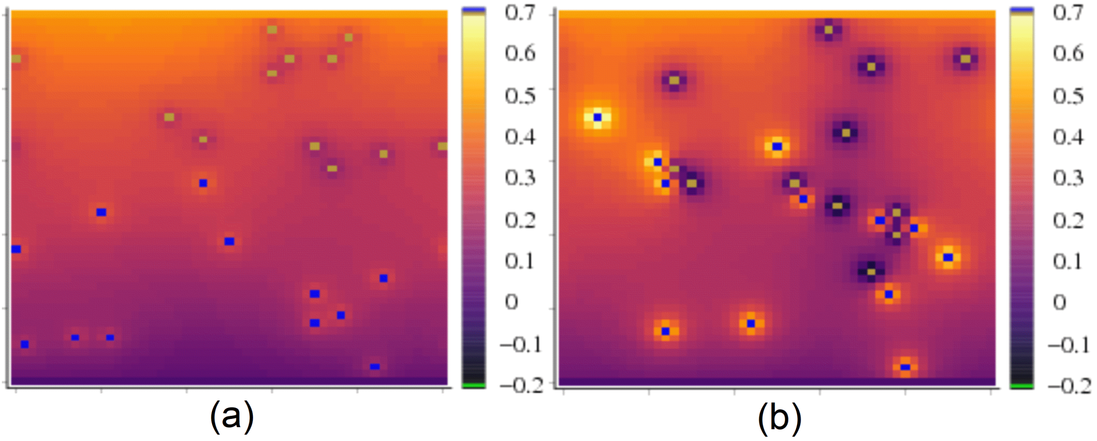


Figure 3.4: Snapshot of the potential map for 1 % salts. The color bars indicate the potential difference from the anode potential. The brighter the color, the larger the potential. The dielectric constants of the electrolytes are (a) $\varepsilon_r = 80$ and (b) $\varepsilon_r = 20$. Blue and yellowish-green particles represent the cations and anions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

and anions will be affected. Then we illustrate the electric potential map for 1 % salt concentration with no dendrite formation in Fig. 3.4. When the dielectric constant is high (Fig. 3.4a), the color gradually becomes brighter from the anode to the cathode, indicating a moderate change in the electric potential between the cathode and anode. The cations (blue) and anions (yellowish-green) are isolated from each other. However, ion pairing and ion clustering occur as the dielectric constant decreases to $\varepsilon_r = 20$ (Fig. 3.4b). The spatial changes in the color gradation between the cathode and anode becomes relatively more inhomogeneous. These preliminary results suggest that both the ion aggregation and the resultant spatial inhomogeneity of the electric potential may become critical in considering the dendrite growth.

3.3.2 Symmetric ions with $d_+ = d_-$

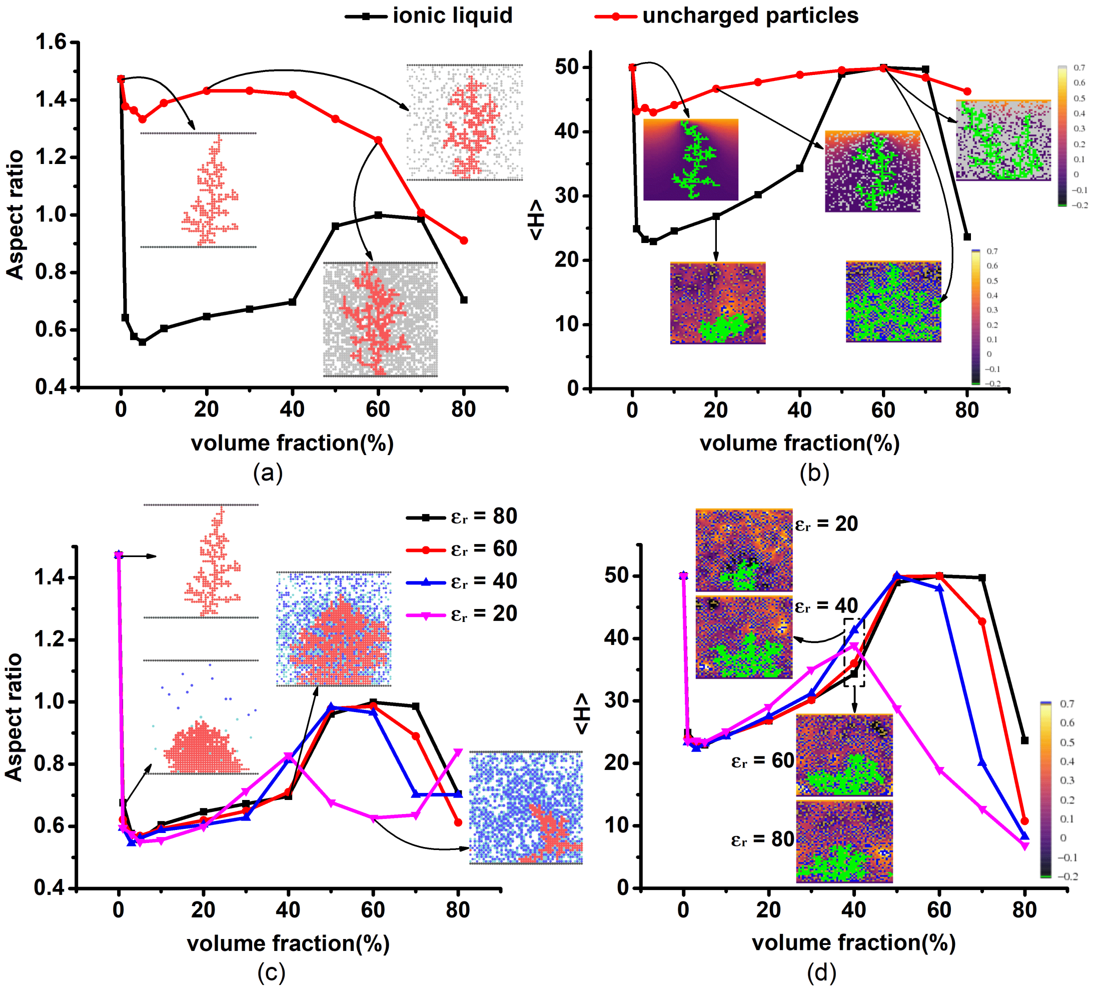


Figure 3.5: Effects of the ionic charge, the dielectric constant on dendrite growth for the symmetric ions ($d_+ = d_- = 2$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

We now consider the dendrite formation (Fig. 3.5). In Fig. 3.5a and Fig. 3.5b, the black line indicates that the addition of tiny amounts of salt ions (≤ 5 %) causes significant decreases in the aspect ratio and average height $\langle H \rangle$. However, the changes in both the aspect ratio and average height $\langle H \rangle$ with increasing salt concentration vary non-monotonically, and thus there appear to be optimal values of salt concentration to maximize the uniformity of the dendrite shape and to inhibit the dendrite growth. Here, we attribute the rise from the local minima around 2 % of the volume fraction to increases in the number of ion pairs that do not participate in electrostatic shielding. Moreover, to consider the significance of the electrostatic interactions for dendrite growth relative to that of the excluded volume interactions, we eliminate the ionic charges on the salt ions. In other words, for comparison, we replace the salt ions with uncharged particles. However, the addition of the uncharged particles (1) does not substantially decrease the aspect ratio unless the volume fraction exceeds about 50 % (Fig. 3.5a) and (2) causes no noteworthy change in the average height $\langle H \rangle$ (Fig. 3.5b). These results suggest that the large inhibition of dendrite growth arises primarily from the effect of the electrostatic interactions.

Importantly, the salt ions cause the spatial gradient of the electric potential to be more locally inhomogeneous but more globally isotropic than the uncharged counterparts do. This is particularly notable in the regions around the dendrite (see the insets in Fig. 3.5b). Here, note the basic rule that the tip of the dendrite tends to grow to the brighter regions for the positive, higher electric potential. However, due to the global uniformity of the potential gradient near the dendrite, such a dendritic formation becomes notably more isotropic. Thus, we suggest that this effect facilitates the lateral growth of the dendrite, significantly flattening the dendrite shape.

Next, we consider the effect of the dielectric constant on the dendrite growth (Fig. 3.5c and Fig. 3.5d). As in Fig. 3.5a and Fig. 3.5b, sharp decreases in the aspect ratio Fig. 3.5c) and average height (Fig. 3.5d) from 0 % to 5 % occur. Here, the difference in the dielectric constant does not notably change the overall trend of the aspect ratio and the average height when the lattice occupancy of the salt ions is $\lesssim$ 40 %. However, when the volume fraction exceeds about 40 %, the effect of the dielectric constant becomes distinct. It is not entirely clear why decreasing the dielectric constant to $\varepsilon_r = 20$ causes the local maximum point to shift to the region at the lower salt concentration. Nevertheless, the potential map in the inset shows that the dendrite in the electrolyte with $\varepsilon_r = 20$ is surrounded by the black regions that indicate that the electric potential is lower than that of the anode. Thus, as indicated by Eq. 2.5, these negative potential spots appear to inhibit the dendrite growth.

Fig. 3.6a and Fig. 3.6b show the effect of the model parameter η in Eq. 2.5 on the dendrite growth. The average height (Fig. 3.6b) is significantly affected by the η value, whereas the changes in aspect ratios are relatively insignificant. Moreover, the difference in the trend of the average height among the lines (Fig. 3.6b) changes non-monotonically. That is, the results change as “black” $\rightarrow$ “red” $\rightarrow$ “blue” $\rightarrow$ “purple” when increasing the η value from 0 to 2. Fig. 3.6b also indicates that the decrease in the average height when increasing the salt concentration is monotonic when $\eta = 0.5$. Thus, in terms of the Li-ion battery design, we do not need to be concerned about

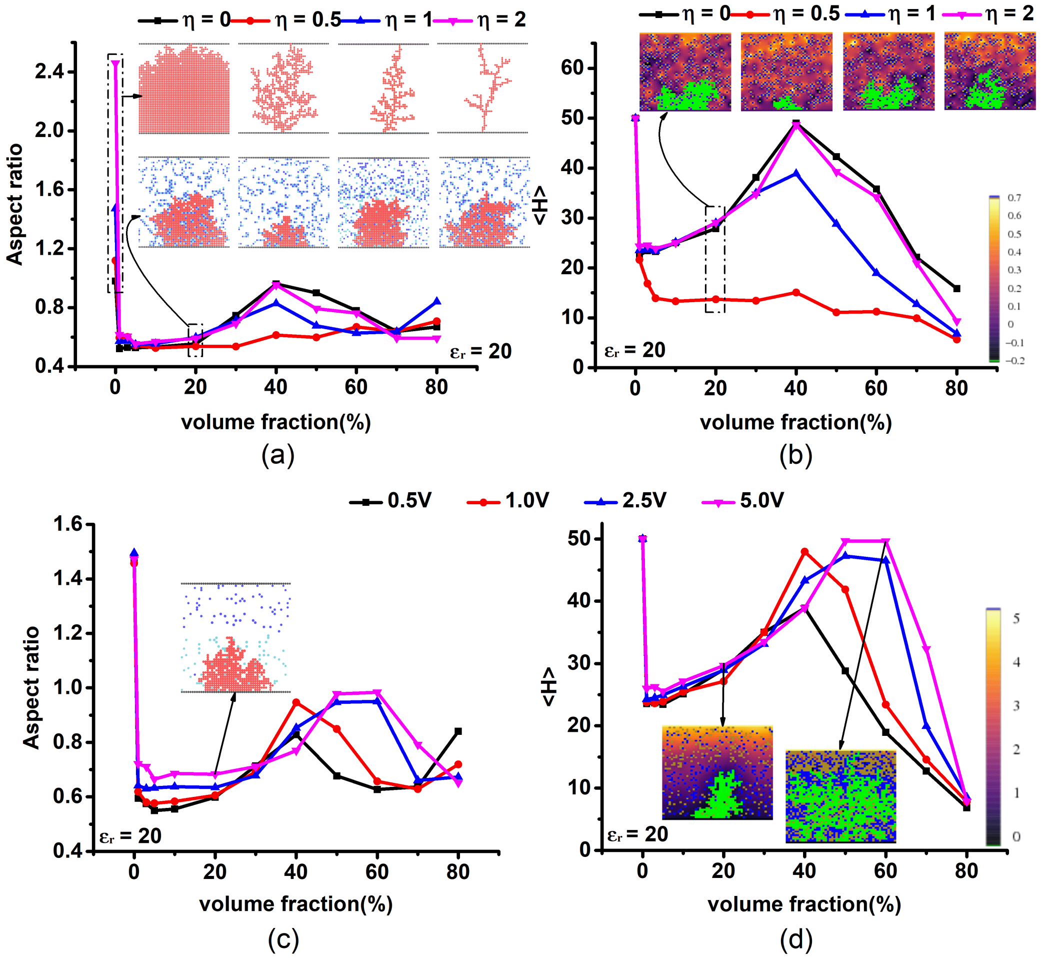


Figure 3.6: Effects of the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($d_+ = d_- = 2$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

the optimal salt concentration that maximally inhibits the dendrite growth.

Fig. 3.7 illustrates the significance of the value η for the dendrite growth at 60 % salt concentration. The results also suggest that it is important to evaluate the η value [or the accurate form of the probability in Eq. 2.5] using more atomistic models that account for the probability of dendrite deposition.

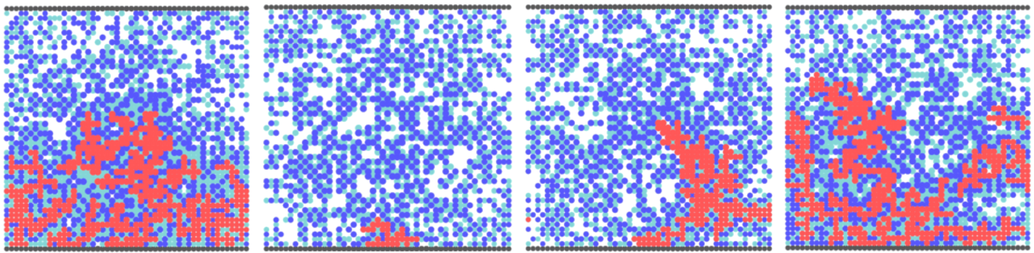

(a) $\eta = 0$ (b) $\eta = 0.5$ (c) $\eta = 1$ (d) $\eta = 2$

Figure 3.7: Snapshots of the dendrite structure with different η values for 60% symmetric ILs and the dielectric constant $\varepsilon_r = 20$. The light blue and blue points correspond to the cations and anions, respectively. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

In Fig. 3.6c and Fig. 3.6d, we change the applied voltage from 0.5 V to 5.0 V. In the case of large voltages such as 5.0 V, we observed the charge separation between the conducting plates, as illustrated by the inset of Fig. 3.6c. The overall trends of both the aspect ratio and average height remain similar when the applied voltage is changed. However, notable non-monotonic variations in these values occur when the salt concentration exceeds about 40 %. Accordingly, 0.5 V becomes the optimal operating voltage to inhibit the average height.

In Fig. 3.8a-d, we also consider relatively larger symmetric ions and their uncharged counterparts with $d = 4$ Å to study the size effect. These large ions also mimic an IL. Interestingly, unlike the result in Fig. 3.5b, the larger uncharged particles promote the dendrite inhibition (Fig. 3.8b), probably due to the larger excluded volume. Nevertheless, the ions more significantly inhibit the dendrite growth. Moreover, unlike the result in Fig. 3.5b, the inhibition trend is nearly monotonic. Therefore, ILs would be more tractable in designing Li-ion batteries.

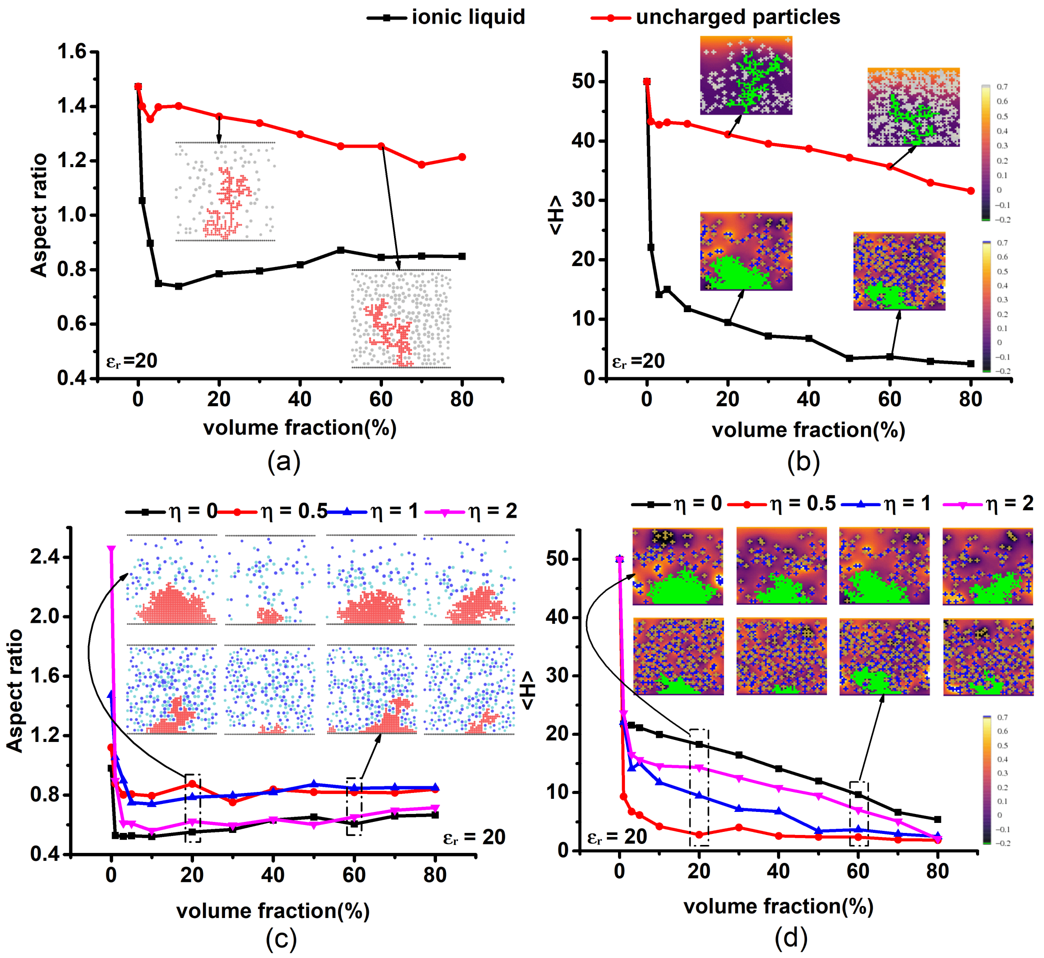


Figure 3.8: Effects of the ionic charge and the parameter η on dendrite growth for the symmetric ions ($d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

Note that changes in the η cause substantial differences in the aspect ratio and average height (Fig. 3.8c and Fig. 3.8d). As in the case of Fig. 3.6b, $\eta = 0.5$ gives rise to a significant inhibition of the average height, and the inhibition effect is even more significant. Note that the overall trends of the aspect ratio and average height are the same among the different η values. Thus, we suggest that a choice of the η value does not alter our conclusion obtained in the present study.

3.3.2.1 Brief summary

To conclude this subsection, we briefly summarize the mechanism of the dendrite growth with the symmetric small ions as follows: (1) Small amounts of salt ions enhance electrostatic shielding near the dendrite and particularly its pointy regions, and accordingly the dendrite growth is significantly inhibited. (2) Increasing the salt concentration enhances ion clustering, and such ionic aggregates reduce the number of ions that participate in electrostatic shielding. Accordingly, ion clustering weakens the driving force for the inhibition of the dendrite growth until a very high salt concentration such as 80 % is reached. Our results and simulation animation suggest that when the salt concentration is very high, small ion clusters tend to aggregate and to form larger clusters, coherently moving over the entire system. These ions do not tightly bind to each other, and the ion clusters do not live as independent charge-neutral objects. Such large clusters widely cover the dendrite surface and yield the effect of electrostatic shielding. (3) Alternatively, the large ions that mimic ILs significantly inhibit ion clustering, and this effect continues to inhibit the dendrite growth more significantly than the small ions do. This result would probably rationalize the remarkably uniform, unconventional dendrite growth observed in our experiment in Section 3.4. Nevertheless, the current observation is on the macroscopic scale, whereas our simulation is on the nanometer scale. Thus, further study to bridge the gap between the length scales is required.

3.3.3 Asymmetric ions with $d_+ = 2d_-$

In this subsection, we demonstrate that the size asymmetry between the cation and anion can be critical to inhibit the dendrite growth. We consider the cation and anion diameters to be 4 Å and 2 Å, respectively.

Both the aspect ratio (Fig. 3.9a) and average height (Fig. 3.9b) decrease upon the addition of small amounts of salt ions. However, unlike the symmetric uncharged particles in Fig. 3.5a, the aspect ratio does not notably change even when the salt concentration is as high as 80 %. Importantly, these variations are monotonically decreasing and nearly reach plateaus. Thus, we do not need to be concerned about identifying the optimal salt concentration to inhibit the dendrite growth. The inhibition trend for the asymmetric ions is relatively simpler to control than for the small symmetric ions. Here, we observed in the snapshots that ion pairing and clustering are relatively weakened, probably because the cations are large. Thus, we attribute the observed plateaus in the aspect ratio and average height to this size effect.

In contrast to the results in Fig. 3.5 and 3.6, we observed nearly monotonically decreasing functions in Fig. 3.9c-d and 3.10. The results are almost unchanged compared to the changes in the dielectric constant, the model parameter η, and the applied voltage, except for Fig. 3.10c. As in the case of the symmetric ions (Fig.

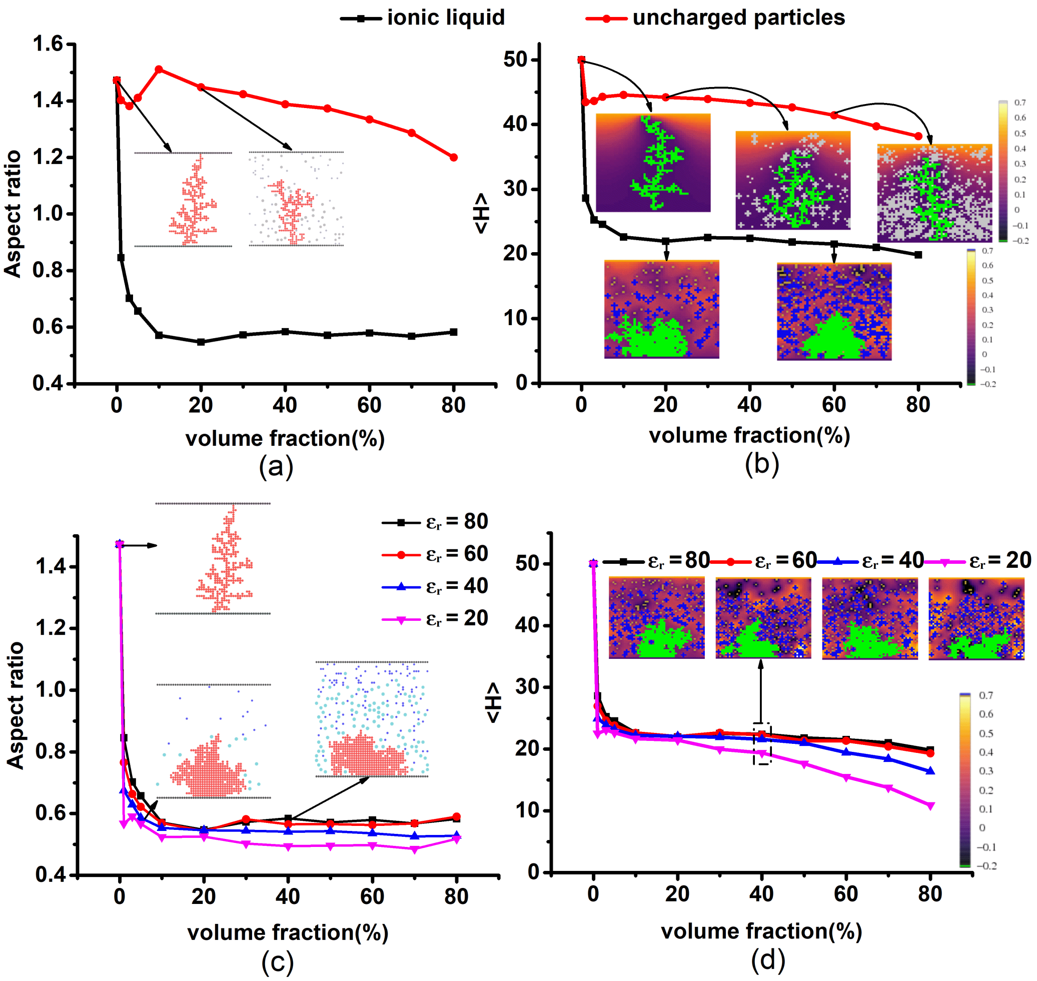
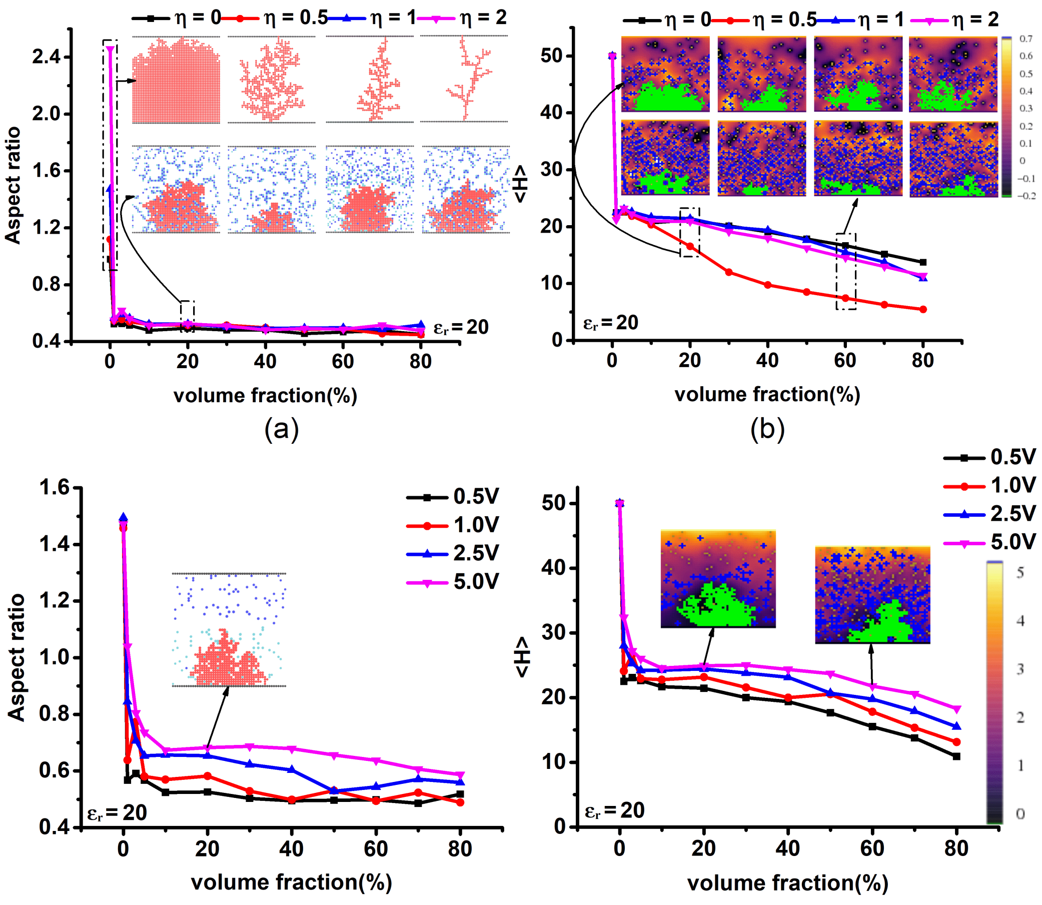


Figure 3.9: Effects of the ionic charge, the dielectric constant on dendrite growth for the asymmetric ions ($d_+ : d_- = 2 : 1$). The x-axis is the lattice occupancy of the salts or uncharged particles. $\langle H \rangle$ designates the average height calculated after $2{\times}10^7$ MC steps. The applied voltage is 0.5 V. The dielectric constants are ε_r= 80. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

3.6), the decrease in the average height is maximized when $\eta = 0.5$ (Fig. 3.10b). Specifically, unlike the symmetric ions, the effect of the applied voltage on the average height $\langle H \rangle$ is relatively insignificant over nearly the entire salt concentrations (Fig. 3.10d). These results also suggest that the inhibition of dendrite growth can be



(a) (b)

(c) (d)

Figure 3.10: Effects of the parameter η and the applied voltage on dendrite growth for the asymmetric ions ($d_+ : d_- = 2 : 1$). The x-axis is the lattice occupancy of the salts or uncharged particles. $\langle H \rangle$ designates the average height calculated after $2{\times}10^7$ MC steps. The applied voltage is 0.5 V in (a)-(b). The dielectric constants are ε_r= 20. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

relatively easier when the cation size is large enough to weaken ion pairing and clustering. Probably due to this fact, ILs would serve as a good substance to substantially inhibit and flatten the dendrite growth.

3.3.4 Asymmetric ions with $2d_+ = d_-$

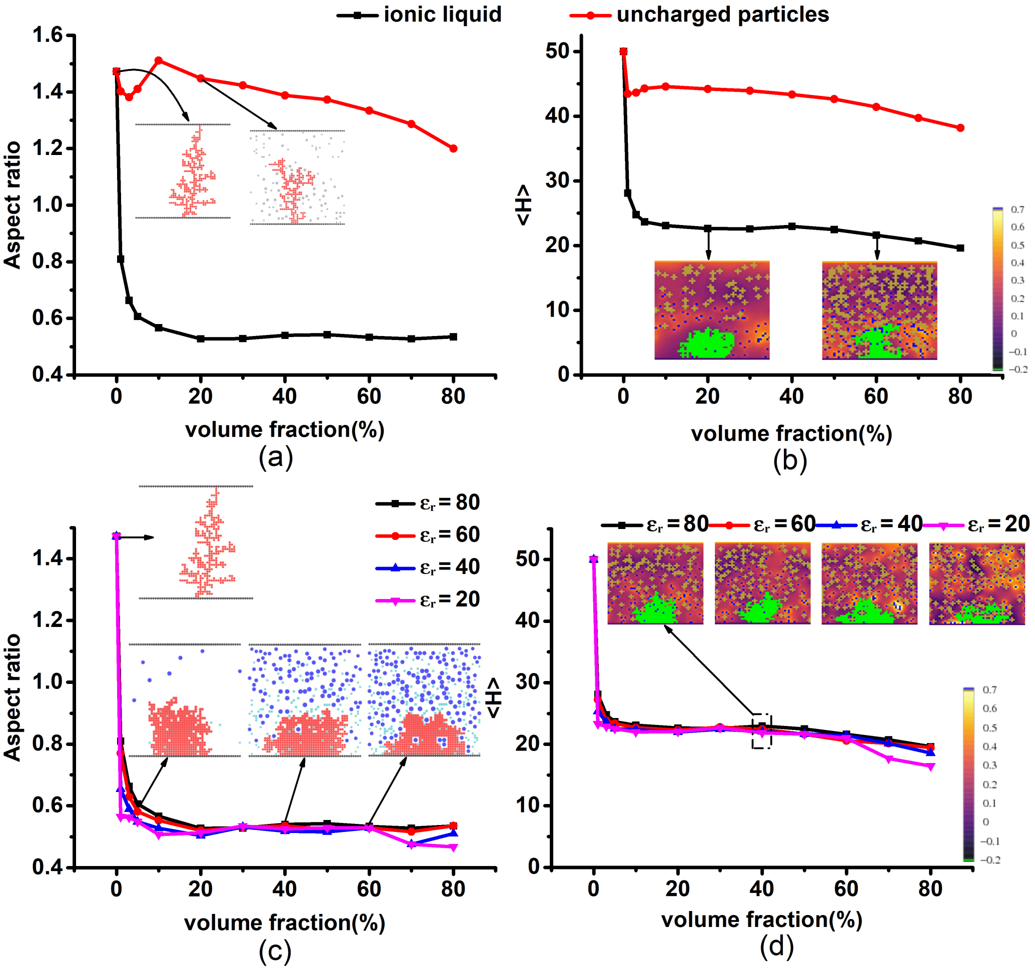


Figure 3.11: Effects of the ionic charge, the dielectric constant, the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($2d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

Finally, we consider the cation and anion diameters to be 2 Å and 4 Å, respectively. The overall trends of the results in Fig. 3.11, Fig. 3.12 are very similar to those in Fig. 3.9, Fig. 3.10. Nevertheless, unlike the results in Fig. 3.9, the effect of the

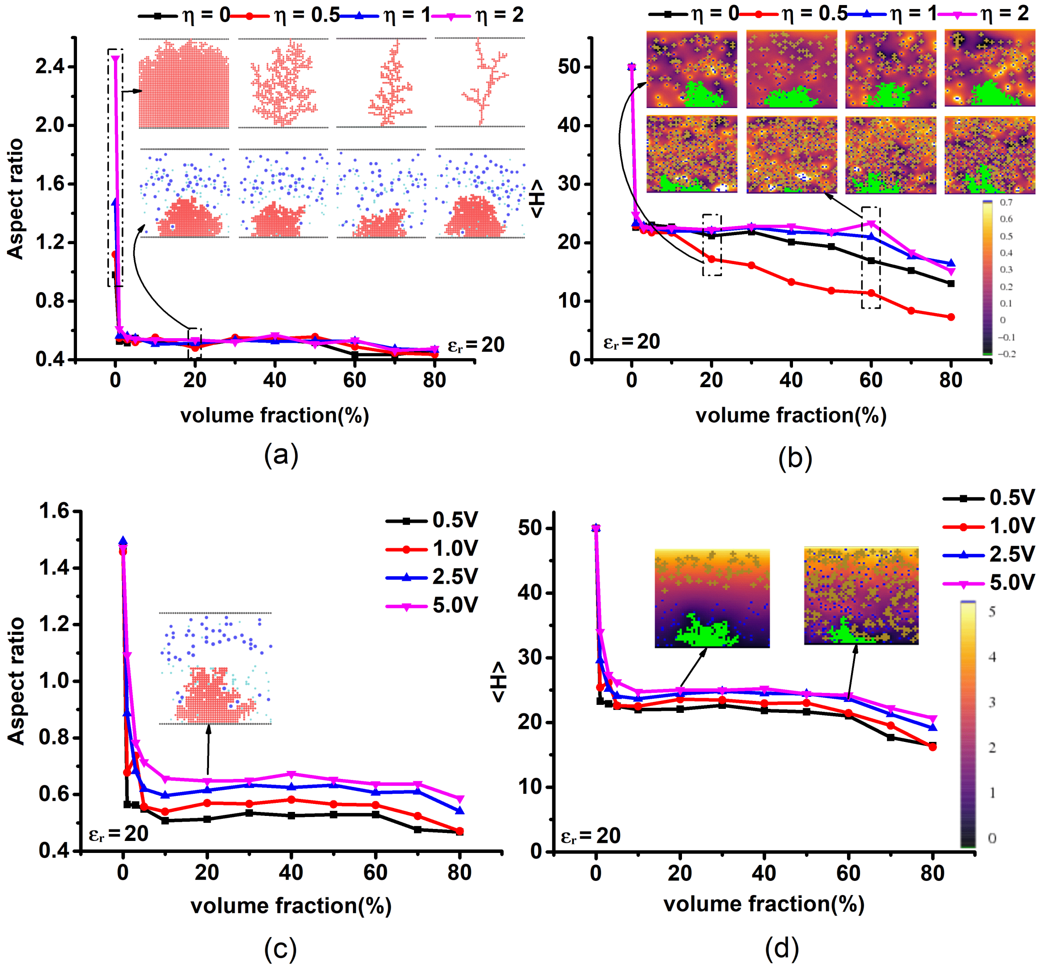


Figure 3.12: Effects of the ionic charge, the dielectric constant, the parameter η, and the applied voltage on dendrite growth for the symmetric ions ($2d_+ = d_- = 4$Å). Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

dielectric constant on the aspect ratio and average height $\langle H \rangle$ becomes insignificant (Fig. 3.11c and Fig. 3.11d). Moreover, the increase in the voltage slightly increases the dendrite growth (Fig. 3.12c and Fig. 3.12d). Thus, we suggest from the strong similarity between Fig. 3.9, 3.10 and Fig. 3.11, 3.12 that a critical factor to inhibit dendrite growth is not to change the type of the size asymmetry in ILs, but instead, to decline ion pairing and clustering.

3.4 Experimental Observation

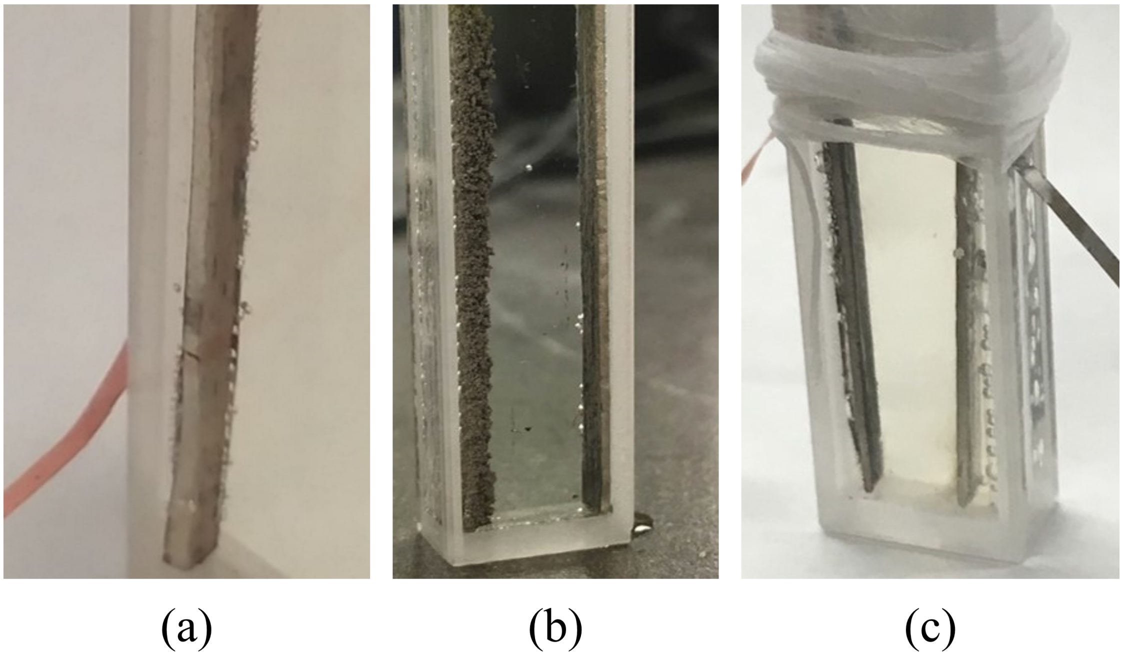


Figure 3.13: Lithium dendrite growth. (a) Original Li foil, (b) lithium deposits under 1M $LiPF_6$ in EC/DMC, and (c) lithium deposits under 1M $LiPF_6$ in EC/DMC + 10% [BMIM][TFSI]. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

We illustrate our experimental observation of lithium dendrite growth in a combination of 1M $LiPF_6$ in EC/DMC electrolyte and 10% IL [BMIM][TFSI]. Fig. 3.13 shows that the addition of the ionic liquid significantly suppresses lithium dendrite and makes the growth trend remarkably uniform in the lateral direction.

3.5 Conclusion

The present lattice model is considered a cellular automaton, in which a minimal assigned rule on a regular grid of cells yields a qualitative feature of physical phenomena. Typically, cellular automata involve model parameters to reproduce a certain degree of complexity in targeted growth mechanisms [77, 78]. Here, note that our present treatment of the electrostatic interactions does not account for the effects of image charges and the diffusion of the lithium ions near the dendrites, whereas these factors may also affect the dendrite growth. However, we assume that the electrostatic interactions due to these effects are electrically screened to a large degree by the highly concentrated ions, and thus the effects are relatively weaker. Therefore, we attribute the two effects to the model parameter η for our cellular automaton model, as the consistency between the present computational algorithm for implicit-metal ion models and the observed fractal dimension of zinc dendrites was demonstrated by Chen and Jorné [24].

We have developed the DLA model proposed by Niemeyer [26] and later developed for zinc dendrite growth by Chen and Jorné [24]. The present study was also motivated by our recent experimental observation regarding remarkably uniform, unconventional lithium dendrite growth in Fig. 3.13. Despite the simple algorithm of the growth patterns, the original toy model accounts for the fractal dimensions observed for dielectric breakdown and zinc dendrite. In other words, the model contains an adjustable parameter that controls the growth probability, but it can be fixed when the fractal dimension is determined in experiments. The present study relies on the strong applicability of the model to the pattern formation due to the electrostatic fields. Thus, we applied the algorithm of the pattern formation to metal dendrite growth in salt-containing liquids between the conducting plates. The main conclusion in our study is summarized as follows: (1) the addition of salt ions can significantly flatten the dendrite shape and inhibit the dendrite growth, primarily due to electrostatic shielding near the pointy regions of the dendrite. (2) It is critical to weaken ion pairing and clustering in electrolytes to inhibit the dendrite growth. Our results show that small salt ions appear to have difficulty achieving this requirement and lead to the undesirable non-monotonic behavior of the dendrite growth as a function of the salt concentration. However, large salts such as ILs can be dissociated near the operating temperatures to a great degree to solve this issue. (3) The size asymmetry that occurs when either the cation or the anion is large (4 Å or larger) affects the dendrite's shape and growth, primarily because ion pairing and clustering are significantly inhibited. In this case, the inhibition of the dendrite growth is not significantly affected by changes in the dielectric constant and applied voltage. The overall conclusion from (1) to (3) remains unchanged when $0 \leq \eta \leq 2$, and appears to hold true in general. Thus, we suggest that ILs be a prospective material to inhibit metal dendrite formation. Among others, the present simulations suggest that ILs with large size asymmetry between the cation and anion serve as a good electrolyte in lithium-ion batteries, as consistent with our experimental observation for lithium deposits under 1 M LiPF6 in EC/DMC + 10% [BMIM][TFSI] in Fig. 3.13.

Concomitantly, the 2D growth appears to be more constrained than the 3D growth. In the case of a 3D model, the dendrite growth in the new lateral direction may easily occur when the ions inhibit the longitudinal growth of the dendrite toward the cathode. Still, this lateral growth appears to facilitate the uniformity of the dendrite surface. Thus, we anticipate that our overall conclusion remains unchanged even in a 3D model.

Chapter 4

Surrogate Model for Prediction of Lithium Dendrite Growth

4.1 Introduction

Nevertheless, the generalization of the MC simulations with various model parameters and the achievement of good statistical convergence with a large number of samples are not easy tasks in general. Moreover, we are aware that like other simulation techniques for dendrite formation, the present lattice MC simulation model may still involve a limitation on the direct comparison between the simulation and experimental data because real systems are typically larger than the simulation length scale, and the entire description of the electrostatic nature on the nanoscales regarding the dendrite formation is often significantly challenging. Thus, we also need a generic, tractable method that provides the design principle of electrochemical devices for the dendrite inhibition in electrolytes. To this end, we also considered a surrogate model using ENNs in machine learning techniques, or their variant, have proved to be useful tools that can outperform a single NN in various studies [85, 86, 87, 88, 89]. We show that our ENNs reasonably captures the non-monotonic trends of the MC simulation results, even though the number of simulation samples at each data point is decreased from the order of 200 to 20. Thus, our surrogate model serves as a convenient tool that enables fast characterization and design of the electrolytes with dendrite formation.

4.2 Computational methods

4.2.1 Data

The data for ENNs are produced from the MC simulations discussed in the previous chapter. The inputs are factors that affect dendrite growth, they are volume fraction φ (%), dielectric constant ε_r, model parameter η, and applied voltage V. The outputs are dendrite height H, and its aspect ratio A. The range of volume fraction is 0.0% - 80.0%, the range of dielectric constant is 20.0 - 80.0, the range of model parameter

Input				Output	
φ	ε_r	η	V	H	A
1.0	20.0	1.0	0.5	23	0.589744
3.0	20.0	1.0	0.5	22	0.536585

Table 4.1
Input variables used in this study

is 0.0 - 2.0, the range of applied voltage is 0.5 - 5.0 V. Because the size of battery system is 51 × 51, the values of dendrite height ranges from 0 - 50, and the values of aspect ratio ranges from 0.0 - 2.0. Based on those, I selected 20 samples from each (φ, ε_r, η, V) parameter space. Examples of the inputs and outputs are shown in Table 4.1.

4.2.2 Design of ENN

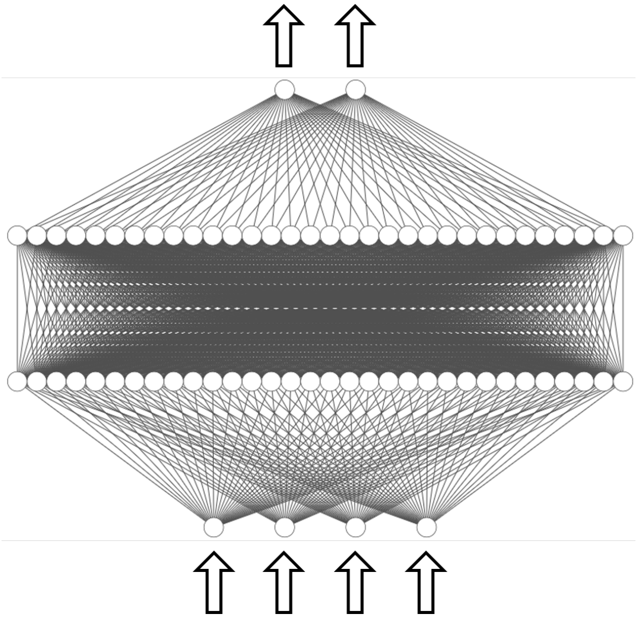
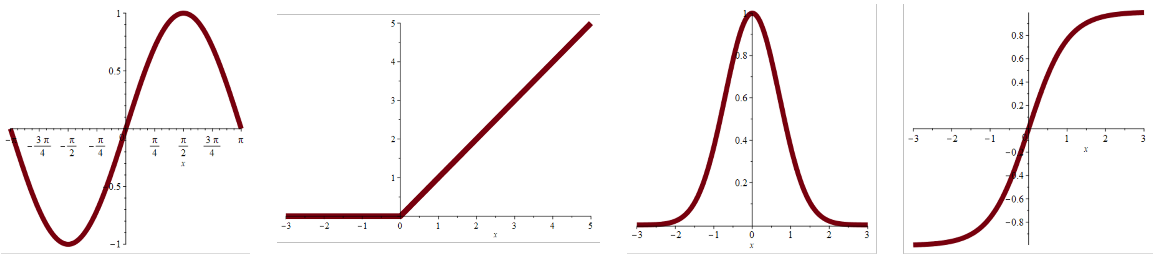


Figure 4.1: Architecture of subNN for lithium dendrite growth prediction.

We employed 20 samples for each volume fraction to train our ENNs using Keras [30]. These samples are split into three parts: training, validation, and test datasets. As illustrated in Fig. 2.6, our ENN consists of an ensemble of 8 sub-NNs (Fig. 4.1) that have identical architecture with different activation functions. They are combined using linear regression. The topology of sub-NNs is 4-32-32-2, which represents the number of neurons in an input layer, the first hidden layer, the second hidden layer, and an output layer, respectively. We evaluated the model by the mean squared error (MSE) of the predicted value ($\hat{y}$) and the target value (y) as Eq.2.7.

The neurons are fully connected between the neighboring layers, and each connection has a weight. For the activation functions, we used ReLU, Gaussian, tanh, and trigonometric functions (Fig. 4.2). To reduce overfitting the model, we performed the early stopping of training with a patience setting of 20 steps. Finally, we construct a surrogate model that accounts for the lithium dendrite growth via training datasets derived from the MC simulations. We employed the dielectric constant ε_r, the model parameter η, the voltage V, and the volume fraction φ for the input variables and



Sine function **ReLU function** **Gaussian function** **Tanh function**

Figure 4.2: Typical types of activation functions.

set the aspect ratio and the average height in the output layer. In the present study, one hidden layer is not sufficient to capture the training datasets, especially for the non-monotonic trend of the results, but two hidden layers are adequate. From the calculation of the MSE, we also empirically found that the sub-NN with 32 neurons per hidden layer serves as a good architecture.

Table 4.2 and Table 4.3 show our ENNs and sub-NNs, respectively. We use different kinds of activation functions for sub-NNs to increase the diversity. The strucuture "ReLU-X-ReLU" means that using ReLU funtion for the 1st hidden layer, X (X = ReLU, sine, cosine, sigmoid, tanh or Gaussian function) for the 2nd hidden layer, and ReLU function for the output layer.

ENN (ID)	Activation Functions in sub-NN
1	ReLU-X-ReLU, X=sine, cosine, sigmoid, Gaussian
2	ReLU-X-ReLU, X=ReLU, cosine, sigmoid, Gaussian
3	ReLU-X-ReLU, X=sine, tanh, sigmoid, Gaussian
4	**ReLU-X-ReLU, X=sine, relu, sigmoid, Gaussian**

Table 4.2
Activation function structures of ENNs.

Sub-NN (ID)	Activation Functions
1	ReLU-X-ReLU, X = sine
2	ReLU-X-ReLU, X = sine
3	ReLU-X-ReLU, X = ReLU
4	ReLU-X-ReLU, X = ReLU
5	ReLU-X-ReLU, X = sigmoid
6	ReLU-X-ReLU, X = sigmoid
7	ReLU-X-ReLU, X = Gaussian
8	ReLU-X-ReLU, X = Gaussian

Table 4.3
Activation function structures of subNNs in No.4 ENN (bold one).

Hyper-parameters that impact the network's design and training are summarized in Table 4.4.

Hyper-parameter	Value or function
Number of subNN in an ENN	8
Number of hidden layers in a subNN	2
Number of activated neurons each hidden layer	32
Activation function in 1st hidden layer	ReLU
Activation function in 2nd hidden layer	{Sine; ReLU; Sigmoid; Gaussian}
Activation function in output layer	ReLU
Patience of early-stopping	10
Train-test split ratio	4:1
Optimization algorithm	Adam optimizer
Learning rate in optimization algorithm	0.001
Cost or loss function each subNN use	Mean square error (Eq. 2.7)
Accuracy metrics	Mean absolute percentage error

Table 4.4
Hyper-parameters and their values or functions in the ENN model.

4.3 Results and discussion

ENN (ID)	Training Error	Validation Error	Test Error
1	13.582 (25.6%)	13.646 (24.8%)	15.323 (29.4%)
2	13.593 (24.6%)	13.965 (24.2%)	14.959 (28.8%)
3	13.356 (24.7%)	13.348 (23.9%)	14.998 (29.0%)
4	**13.573 (25.5%)**	**13.863 (26.0%)**	**14.823 (29.0%)**

Table 4.5
Training, validation, and test errors of different ENNs. The MAPE is shown in parentheses (Lower is better.)

We trained 4 ENNs with different activation functions (Table 4.2). Table 4.3 shows architectures of sub-NNs of No.4 ENN in Table 4.2. Here, the chance of overfitting tends to decrease as the difference between the training error and validation error

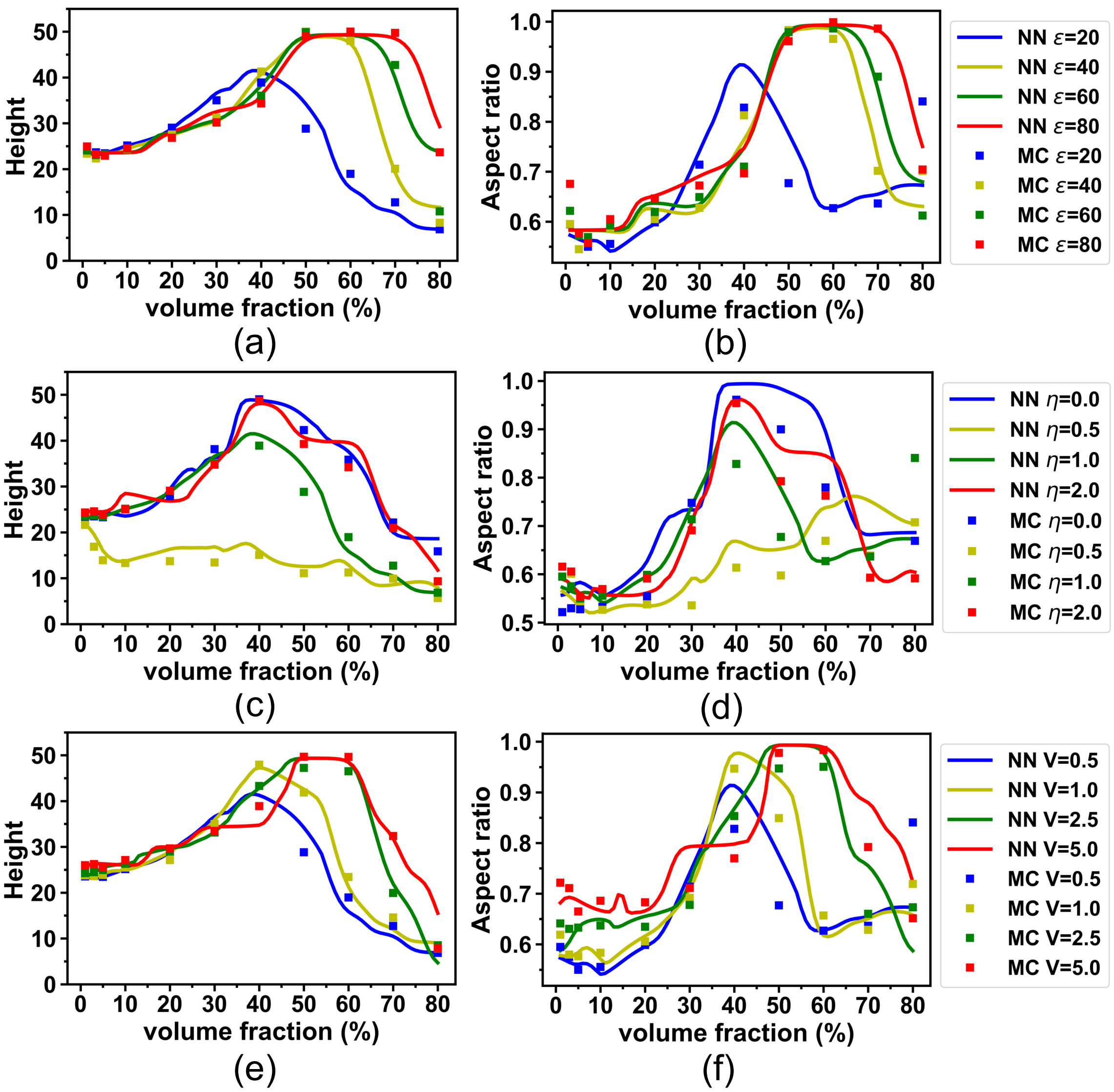
sub-NN (ID)	Training Error	Validation Error	Test Error
1	506.762 (60.8%)	486.218 (60.8%)	495.212 (61.0%)
2	506.761 (59.9%)	486.217 (60.0%)	495.210 (60.2%)
3	507.018 (100%)	486.466 (100%)	495.465 (100%)
4	17.232 (25.0%)	18.275 (26.9%)	18.838 (27.7%)
5	14.942 (26.0%)	15.282 (28.0%)	16.785 (28.9%)
6	14.520 (24.0%)	15.728 (25.8%)	16.814 (26.7%)
7	15.175 (23.9%)	15.658 (25.9%)	17.162 (26.6%)
8	15.528 (25.2%)	15.524 (26.8%)	17.186 (27.5%)

Table 4.6

Training, validation, and test errors of the sub-NNs in No.4 ENN (bold one in Table 4.5). The MAPE is shown in parentheses.

decreases. Note that some sub-NNs are poorly trained with an error of about 500, yet the corresponding ENN outperforms any sub-NNs and thus provides a good prediction with relatively small errors. Thus, the ENNs can efficiently avoid the local minimum in the MSE that is more likely to occur in a single NN.

We plotted the predicted average heights and the aspect ratio of lithium dendrites in Fig. 4.3 and Fig. 4.4. In Fig. 4.3, we compared the MC simulation result with one of the excellent sub-NNs (No.5 sub-NN in Table 4.6) containing the sigmoid function for the activation function in the second layer. Note that we used only 20 samples for each volume fraction (i.e., each data point in the figures) to train each sub-NN. Still, our ENNs can reasonably capture and predict the non-monotonic trend of both the average height and the aspect ratio. Fig. 4.4 shows the comparison between the MC simulation result with the prediction from No. 4 ENN in Table 4.5. Other ENNs also provide similar consistencies between the MC simulation and ENN prediction. Although some sub-NNs may have large errors, their combination with good sub-NNs



(a) (b)

(c) (d)

(e) (f)

Figure 4.3: Comparison between the MC simulation and No.5 sub-NN in Table 4.6 for the average height and the aspect ratio. The solid lines indicate the prediction from the sub-NN in the range of 1% to 80% volume fraction of salts. The squares indicate the statistical averages calculated from the same datasets of the MC simulations as those used in Fig. 3.3c-d and Fig. 3.4a-d. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

tends to provide a robust, relatively accurate ENN. Overall, we suggest that although a single NN can be employed as a surrogate model for the MC simulations, the ENNs

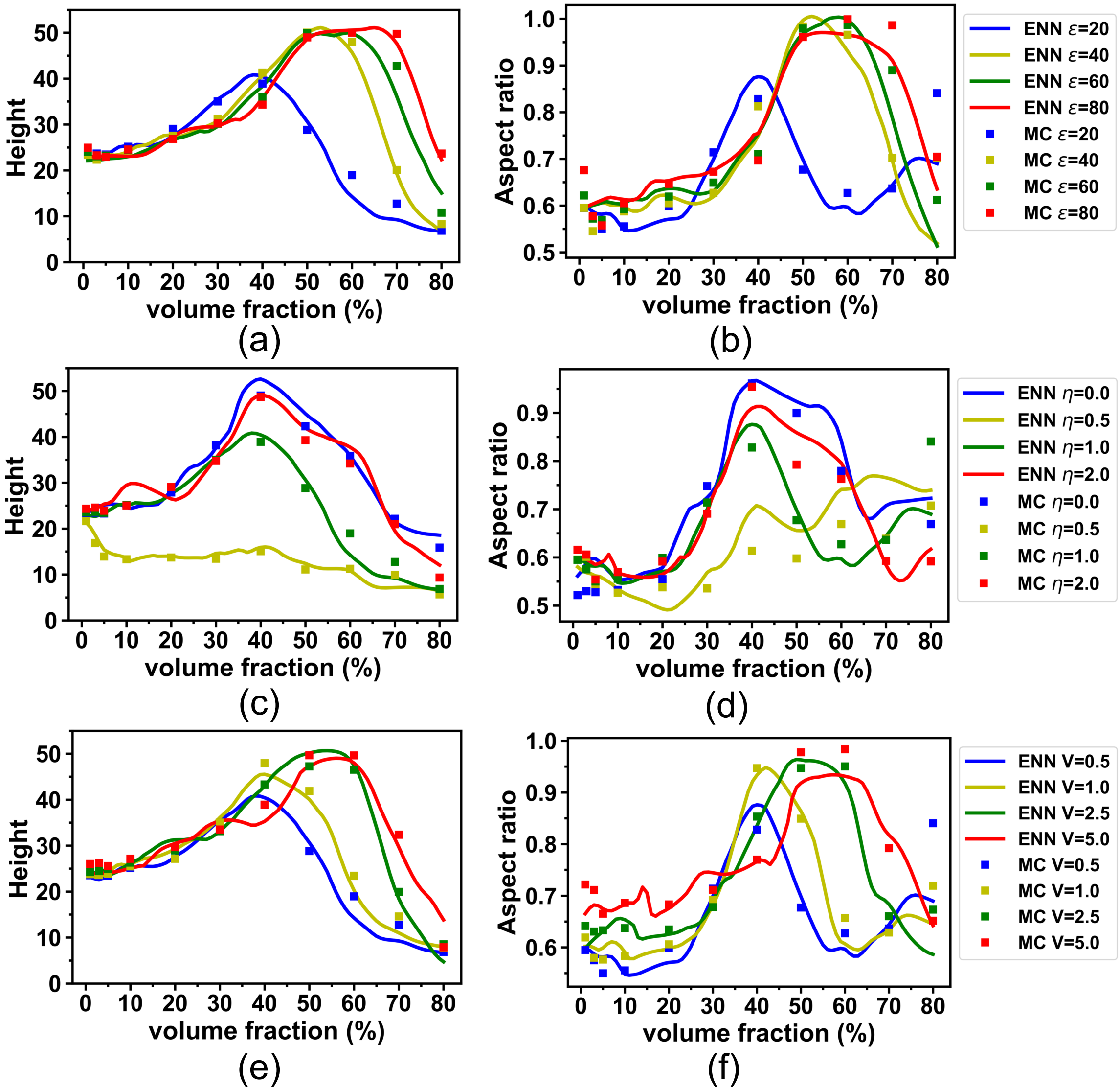


Figure 4.4: Comparison between the MC simulation and ENN prediction for the average height and the aspect ratio. The symbols and lines are defined as in Fig. 4.3. Reproduced from Ref. [1] with permission from the Royal Society of Chemistry.

are not sensitive to initial weights and are robust against the choice of the activation functions, compared to a single NN.

4.4 Conclusion

We considered a surrogate model that accounts for the MC simulations using ENNs. In our study, we found that a single NN can be consistent with the MC simulation results, but ENNs with sub-NNs can be more robust against the choice of the activation functions and can easily be trained from various initial weights. This is mainly because the ENNs can be trained well when poor sub-NNs are combined with good sub-NNs. Importantly, unlike 200 samples for the average of the MC simulation results for each volume fraction (i.e., each data point), we used only 20 samples for training. Thus, our surrogate model serves as an alternative to the computationally demanding MC simulations.

Chapter 5

Ensemble Neural Network for Predicting Dielectric Constants of Pure Solvents

5.1 Introduction

The use of machine learning methods enable the construction of surrogate models that capture highly nonlinear features that are dependent on a number of model parameters, which used to be a challenging task when performed using traditional optimization algorithms.

Of particular importance in materials science is that NNs can considerably reduce the required computational expense and burden when applied to molecular dynamics (MD) simulations, as largely reviewed in Ref. [90]. Recent examples of the application of NNs in MD simulations include NNs trained for the free energy of coarse-grained molecules [91], NNs trained on quantum molecular dynamics configurations for the dielectric constant of water [92], ML potentials obtained from NNs for the structure and permeability of water [93], NNs used to produce a Hamiltonian for an electronic structure [94], and an NN-assisted MD method to reduce the computations required when conducting open-boundary simulations [95].

On the technical side of optimization algorithms, an ensemble of linear gross domestic product estimates was shown to enhance the model accuracy and improve its generalization [96]. Along similar lines of research, NN ensembles appear to be straightforward methods for enhancing the versatility of NNs [97] and accordingly improving the accuracy of the resulting models when combined with the optimal linear combination of trained NNs [98]. Thus, the effectiveness of ensemble NNs (ENNs) in various systems has been explored for several decades, and the applications involving MD simulations include ENNs for determining the potential energy surface of a bimolecular reaction [99], a combination of different ML models for studying the dynamics of water droplets [100], and the analysis of the potential energy surfaces of various Lennard–Jones fluids with two types of ENNs [101]. Nevertheless, the study of ENNs combined with MD simulations remains significantly limited, particularly with respect to the dielectric properties of materials [92].

Although the coarse-grained simulations for Stockmayer Fluids (SFs) can be substantially faster than the atomistic simulations [102], the exploration of the molecular parameters that simultaneously account for various salt-free and/or salt-doped liquids over vast parameter space may easily become impractical, especially for multicomponent liquid and ion mixtures.

Thus, a more computationally fast surrogate model that captures the present MD simulation should be critical, at least, in the following two perspectives: First, such a surrogate model considerably facilitates researchers to identify appropriate molecular parameters of SFs in the system involving various types of liquids and, if any, salt ions.

Second, the calculation of the dielectric constant is often statistically noisy due to the thermal fluctuation of molecular polarization, and thus, the large statistical fluctuation tends to make the statistical convergence of the dielectric constant calculation relatively slow. The iterative reparameterization of the model by repeatedly checking the statistics of the dielectric constant after the equilibration for the timescale of 1-20 ns is undoubtedly a daunting task. In this regard, the surrogate model is preferred to be constructed from a relatively small amount of simulation data.

In this chapter and the next chapter, we show that ENNs serve as efficient surrogate models for Stockmayer Fluid MD simulations that provide the dielectric constants of polar solvents and NaCl solutions. These two chapters aim to show how NNs can help address the problem of calculating dielectric constants, and discusses our general theoretical framework for reducing computational costs arising from the statistical ensemble of MD simulations. ENNs trained with MD simulation output datasets, including small datasets obtained with a low-sampling rate, can provide predictions of the dielectric constants for various model parameters without running additional MD simulations. Advantages of the use of ENNs include the following objectives: (1) ENNs can provide the model parameters for multiple solvents and salt ions. (2) Unlike statistical averages calculated from ensembles of vast numbers of configurations, ENNs require relatively small amounts of simulation data for training. (3) Even if a single large NN performed equivalent computations, training data might be too large for the memory. With ENNs, however, various batch and ensemble methods can be used to split the data into parts that can be more quickly processed. (4) Typically, the objective (loss) function of NNs has many local minima, and the backpropagation algorithm for the optimization of the objective function often results in a movement towards these minima (the so-called "local minimum problem"). However, ENNs tend to prevent model training from getting stuck in these local minima.

5.2 Computational methods

Accordingly, we employ ENNs in machine-learning methods because this modeling worked efficiently in the study of lithium dendrite growth in ILs that involved similar computational requirements in Chapter 4. In this chapter, we develop an ENN to predict the dielectric constant of salt-free solvents trained by MD simulation results. Here, an ENN consists of multiple sub-NNs and each sub-NN has its own structure. The idea of this method arises from an analogue of wisdom of crowds, in which a hybrid of all predictions from the sub-NNs should provide a better prediction than the individual.

5.2.1 Data

I collected our in-house datasets regarding averaged dielectric constants from thousands of MD simulations of 7 types of pure solvents, water, MeOH, EtOH, Acetone, 1-Propanol, DMSO, and DMF. They were calculated using MD simulations based on their properties. Those properties are inputs for training, they are dipole moment μ, diameter of solvent molecule σ, number density ρ, and Lennard-Jones (LJ) parameter ϵ_{LJ} for the solvent-solvent interaction. The output is dielectric constants ε_r of pure solvents.

For inputs, the range of dipole moment is 0.0 - 3.7 Debye, the range of diameter is 2.5 - 5.2 Å, the range of LJ parameter is 0.01 - 0.10 kcal/mol, the number density is 0.007 - 0.034 Å^{-3}. For outputs, the range of dielectric constants of them are 0.0 - 212.2. Examples of the inputs and calculated outputs, and their corresponding type of pure solvents are shown in Table 5.1.

Type	Input				Output
	μ	σ	ρ	ϵ_{LJ}	ε_r
Water	1.73	2.5	0.033474	0.01	94.87
MeOH	2.02	4.0	0.014889	0.01	23.64
EtOH	2.13	3.6	0.011333	0.05	15.64
Acetone	2.66	4.6	0.008148	0.05	21.27
1-Propanol	2.7	5.2	0.008037	0.10	32.63
DMSO	3.0	4.3	0.008481	0.10	38.85
DMF	3.0	5.0	0.007815	0.10	49.21

Table 5.1

Examples for input variables and output for pure solvents.

5.2.2 Design of sub-NN

In general, ENN can be considered a learning technique where many neural network models are joined to solve one problem. An ENN is composed of n sub-NNs, and each sub-NN can have its own structure (see Fig. 5.1). By combining all predictions of sub-NNs, like ensemble of wisdom, the ENN can make a better prediction.

We consider an ensemble of 8 sub-NNs that have identical architecture. The topology of sub-NNs is 4 inputs, 11 hidden layers with 30 neurons per hidden layer, and 1 output

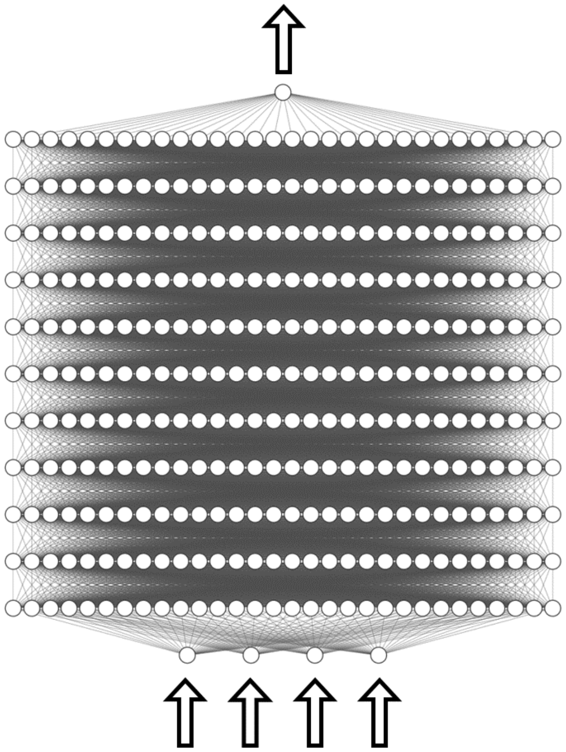


Figure 5.1: The sub-NN structure for the pure solvents. The inputs are the dipole moment μ, the solvent diameter σ, the number density ρ, and the LJ parameter ϵ_{LJ} for the solvent-solvent interaction. The output is dielectric constant ε_r.

(i.e., 4-30-30-30-30-30-30-30-30-30-30-30-1), see Fig. 5.1(a)). We also evaluated the model by MSE (Eq. 2.7).

The weights of all sub-NNs are randomly initialized and are optimized by minimizing the loss function. The MSE indicates the standard error for the evaluation of NN's performance. However, we also provide the mean absolute percentage error (MAPE) $\frac{100\%}{n}\sum|\frac{y-\hat{y}}{\hat{y}}|$ that serves as an alternative measure to analyze the performance. For activation functions, we employ ReLU, Gaussian, and sine functions. Specifically, we found that the sandwich structure such as "X-Y-X" and "X-Z-X", with X =

ReLU is critical for substantially enhancing the performance of the sub-NNs and accordingly the ENN. Specifically, Y = sine, and Z = Gauss. This architecture is the best architecture we can find. To reduce overfitting of the models, we also perform the early stopping of training with a patience setting of 30 epochs. Hyper-parameters are summarized in Table 5.2.

Hyper-parameter	Value or function
Number of subNN in an ENN	4
Number of hidden layers in a subNN	11
Number of activated neurons each hidden layer	30
Activation function for hidden layers	X-Y-X-Z-X-Y-X-Z-X-Y-X, where X=ReLU, Y=sine, and Z=Gauss
Activation function in output layer	ReLU
Patience of early-stopping	30
Batch size	20
Train-test split ratio	4:1
Optimization algorithm	Adam optimizer
Learning rate in optimization algorithm	0.001
Cost or loss function each subNN use	Mean square error (Eq. 2.7)
Accuracy metrics	Mean absolute percentage error

Table 5.2
Hyper-parameters and their values or functions in the ENN model.

We produce 76 samples for the set of polar solvents by changing the dipole moment μ, the solvent diameters σ, the LJ parameters ϵ_{LJ} , and the number density ρ and calculating the dielectric constant ε_r for each sample from the average of 10000 MD configurations. We thus train the ENN with 76 averaged values using the Keras API [30].

To enhance the model accuracy, we duplicated these 76 samples 5 times to employ

76×5 samples and divided all data into two parts: training datasets and validation datasets. Normally, data samples are classified into training, validation, and test datasets. However, we only split them into training and validation datasets to maximally utilize the limited number of samples for training. After training all sub-NNs, outputs from them are combined to yield the output of the ENN via linear regression.

5.3 Results and discussion

Pure Solvents	Training Error	Validation Error
Sub-NN1	29.1 (17.2%)	25.5 (15.7%)
Sub-NN2	20.3 (20.3%)	18.1 (19.0%)
Sub-NN3	11.8 (7.2%)	8.2 (7.0%)
Sub-NN4	26.9 (21.6%)	29.6 (20.4%)

Table 5.3
Training and validation errors (MSEs) of sub-NNs for pure solvents. The MAPE is shown in parentheses. The topology of the sub-NNs is 4-30-30-30-30-30-30-30-30-30-30-30-1, and the sequence of the activation functions consists of X-Y-X-Z-X-Y-X-Z-X-Y-X-X, where X=ReLU, Y=sine, and Z=Gauss.

Pure solvents	Constituent sub-NNs	Training Error	Validation Error
ENN1	**1,2,3**	**11.7 (6.3%)**	**7.9 (8.3%)**
ENN2	1,2	18.6 (22.2%)	15.0 (26.2%)
ENN3	1,2,4	16.4 (15.0%)	14.1 (17.0%)

Table 5.4
Training and validation errors (MSEs) of different ENNs for salt-free solvents. The MAPE is shown in parentheses.

Table 5.3 and 5.4 show training and validation errors of ENNs and sub-NNs, respectively. Sub-NNs have the same topology and activation function, the difference among them is the initial weights. ENNs in Table 5.4 are combinations of sub-NNs in Table 5.3. ENN1 consists of sub-NN1 and sub-NN2. ENN2 consists of sub-NN1, sub-NN2, and sub-NN3. ENN3 consists of sub-NN1, sub-NN2, and sub-NN4. Note that most sub-NNs are poorly trained with an error of 18 - 30, among them, we can find that No.3 sub-NN is the best subNN. By including the best sub-NN in ENN, the ENN achieves the very low error. And we can also conclude that the corresponding ENN outperforms any sub-NNs and thus provides a good prediction with relative small errors. Unlike the sub-NNs, the errors of ENN1 are not substantially changed by the initial model parameters. The robustness of the training performance to the combination of sub-NNs arises mainly from the fact that in linear regression, well-trained sub-NNs are weighted more significantly than sub- NNs with relatively poor performance.

Because the numbers of training data and validation data are limited, one or two poor prediction may cause a significant increase on MAPE, causing a “poor-trained” subNN. Thus, the ENNs can efficiently avoid the local minimum in the MSE that is more likely to occur in single NN.

We plotted the dielectric constant of 7 solvents with different LJ parameters in Fig. 5.2. We focus on the solvent diameter and the LJ parameter because the Onsager equation (the standard mean-field theory for the dielectric constant) cannot predict changes in the dielectric constant with respect to the two [103]. Here, the total

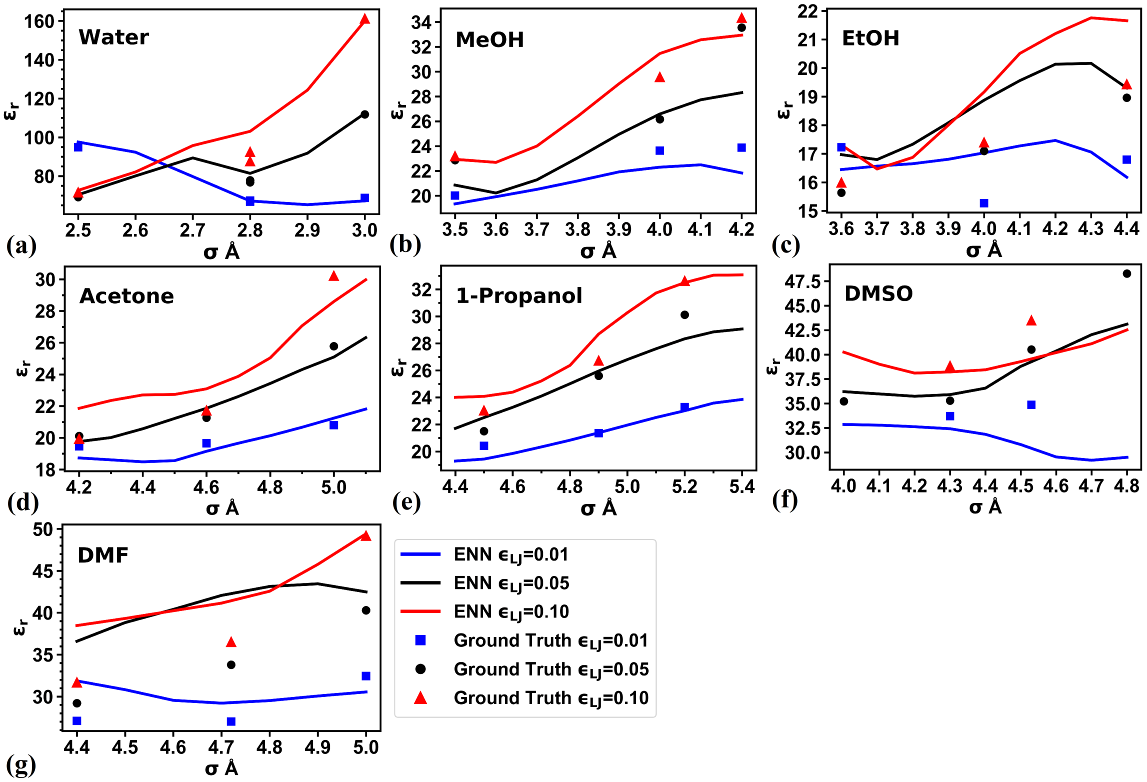


Figure 5.2: Dielectric constants predicted by No.1 ENN for different pure solvents with ϵ_{LJ}=0.01 (squares), 0.05 (circles), and 0.10 (triangles). The solid lines indicate the prediction from the ENN, solid symbols indicate the statistical averages calculated from the MD simulations. Solvents are: (a)water, (b)MeOH, (c)EtOH, (d)acetone, (e) 1-Propanole, (f)DMSO, and (g)DMF. Reproduced from Ref. [2] with permission from the Materials Research Society.

number of training data are limited to 76 samples for all the solvents, and Fig. 5.2 includes 65 samples.

In Fig. 5.2, we compared the MD simulation result with No.4 ENN prediction. Note that we use 1 sample for each diameter (1 sample for each point in Fig. 5.2) to train each sub-NN, however, the sample is obtained from 10,000 MD simulations, so that the training sample is the “true” statistical average compared to one MD simulation sample. Overall, the ENN reproduces the trends in ε_r for almost all of the solvents, and is sufficient to guide the selection of model parameters to obtain a particular value of ε_r (e.g., for future MD simulations). Here, we note that conventional NNs consisting of only ReLU and sigmoid functions are unlikely to capture the observed variations in the dielectric constants, such as the nonmonotonicity between the solvent diameter σ and the LJ parameter ϵ_{LJ}. To address this nonmonotonicity, we find that the inclusion of a sine function sandwiched by ReLUs considerably improves the training efficiency, and a Gaussian function sandwiched by ReLUs further improves the results. As discussed before, however, the combination of these sub-NNs is the essence of the substantial improvement and robustness of the model accuracy, mitigating the local minimum issue and providing the model robustness against poorly trained sub-NNs. Concomitantly, we test BN in one or more layers and find that BN makes the prediction results worse. The reason for this could be due to the insufficient amount of training data. Nevertheless, in the case of NaCl solutions in the next chapter, we show that BN averages the results and improves the performance of ENNs when the training data are statistically noisy.

5.3.1 Reorganizing ENN with 3 subNNs

Since No.3 subNN is the best subNN, and No.1, No.2 and No.4 are almost the same level training, we ensemble these three "poor-trained" sub-NNs together to make a prediction. Results are shown in Fig. 5.3. Comparing with Fig. 5.2, we can see that three bad-trained sub-NNs can also make a good prediction, even though the predictions in Fig. 5.3 are not as good as the predictions made in Fig. 5.2. For EtOH, sub-NN3 captures its features, so that ENN including sub-NN3 can predict the EtOH's dielectric constant better.

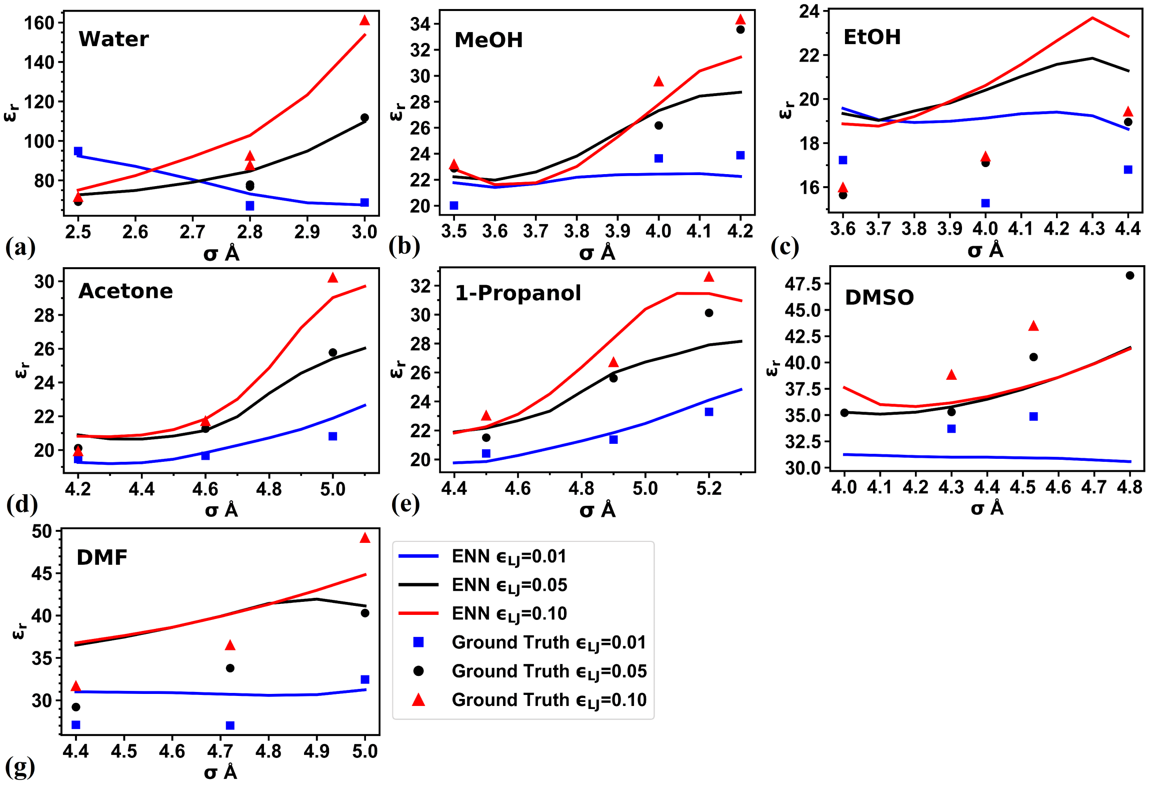


Figure 5.3: Dielectric constants predicted by No.3 ENN consisting of three poorly trained sub-NNs (sub-NN1, sub-NN2, and sub-NN4). The solid lines indicate the prediction from the ENN, solid symbols indicate the statistical averages calculated from the MD simulations. Solvents are: (a)water, (b)MeOH, (c)EtOH, (d)acetone, (e)1-Propanole, (f)DMSO, and (g)DMF.

5.3.2 Effect of combination method

Fig. 5.4 represents the results of a hybrid of three sub-NNs via a combination neural network (com-NN). Outputs from the three sub-NNs become inputs for the com-NN with one hidden layer consisting of 16 neurons. All activation functions in the hidden layer are ReLUs. The change in the number of the hidden layer did not provide any notable changes in the results.

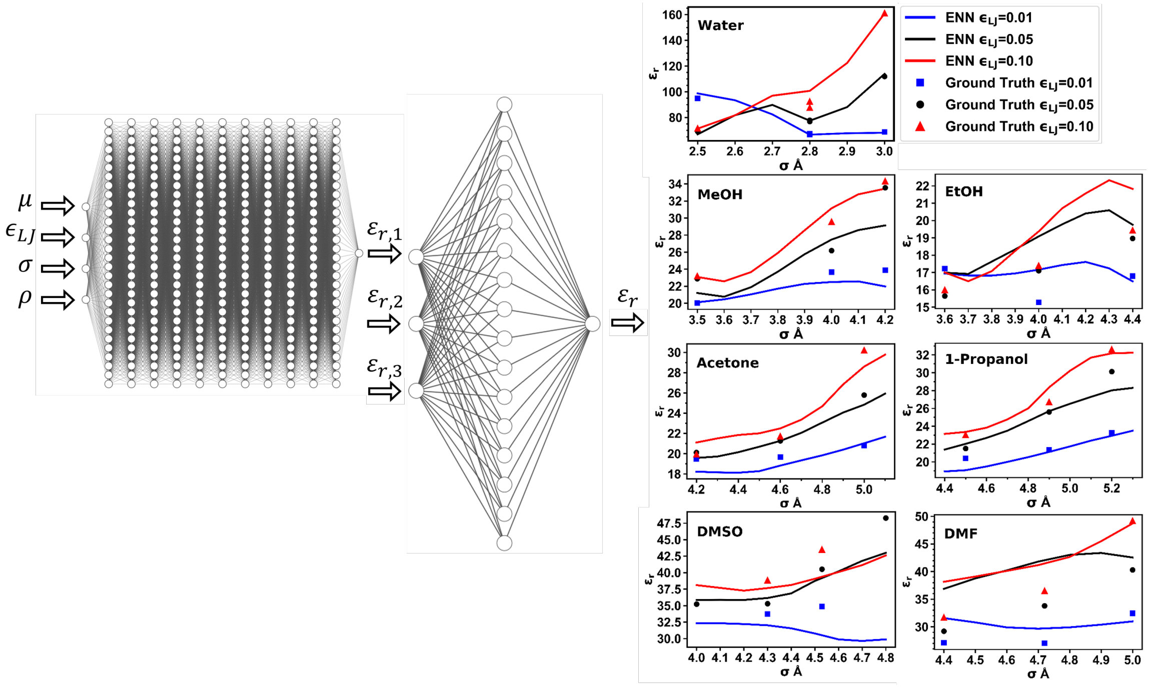


Figure 5.4: Dielectric constants of salt-free solvents predicted by ENN1 consisting of four sub-NNs combined through a neural network. Reproduced from Ref. [2] with permission from the Materials Research Society.

5.4 Conclusion

We constructed ENNs for the calculation of the dielectric constants of polar solvents. For the training data, we used Stockmayer fluid MD simulation methods that accounted for the dielectric responses of dipolar salt-free solvents. We require sine function in the present NN architecture because our ENNs with activation layers containing sine and Gaussian functions captured the trends of the dielectric constants of seven polar solvents. Sinusoidal activation functions often yield large numbers of shallow local minima [104], but our ENNs appeared to mitigate this issue.

Therefore, ENN is more robust to be employed than a single NN as a surrogate model for the MD simulations. Especially, periodic functions like sine functions can be applied in a very deep neural network with limited amount of training data to learn better. And ENNs with well-designed sub-NNs can still provide a steady and reasonable prediction without affected by one or several poor-trained sub-NNs. In addition, the results of the com-NN resemble those derived from the linear regression method. Thus, we suggest that the complexity of the hybridization for the output layer is unlikely to substantially improve the prediction.

Chapter 6

Ensemble Neural Network for Predicting Dielectric Constants of Salt-doped Solvents

6.1 Introduction

In this chapter, an ENN is developed to predict the dielectric constant of salt-doped solvents trained by MD simulation data points. This chapter is organized as follows. Firstly, the data collection and the design of sub-NNs of ENNs will be described in section 6.2. Secondly, results and discussions in terms of the impact of bagging, BN, and the combination method on predictions are presented in section 6.3. Thirdly, a brief conclusion is shown in section 6.4.

6.2 Computational methods

6.2.1 Data

Input		Output
c_s	ϵ_{LJ}	ε_r
0.17931	0.01	28.69
0.53793	0.01	20.55
1.0759	0.10	50.00
1.6138	0.10	12.14
2.6896	0.50	34.28
3.9448	0.50	16.34

Table 6.1
Examples for input variables and output for salt-doped solvents.

We collected every single dielectric constants from 2100 MD simulations of salt-doped solvents with different salt concentrations. We employ the dipole moment 1.73 D throughout all these simulations because this dipole moment provides the dielectric constant 80.2 consistent with the experimental value. Two properties are inputs for training, they are Lennard-Jones (LJ) parameter ϵ_{LJ} for the ion-ion interaction and salt concentration c_s. The output is dielectric constants ε_r of salt-doped solvents.

For inputs, the range of salt concentration is 0.0-5.0 M, the range of LJ parameter is 0.01-0.50 kcal/mol where the upper limit is larger than salt-free solvents. For outputs, the range of dielectric constants of them are 7.5-362.8. Examples of the inputs and calculated outputs, and their corresponding type of pure solvents are shown in Table 6.1. For training data, we randomly select 40 configurations from 2100 configurations collected from simulation runs performed with given salt concentration and LJ parameter.

Thus, the total numbers of the simulation data and the training data amount to 2100×(the number of data points) and 40×(the number of data points), respectively. Here, we represent the latter data set as a poor training data set. To examine the robustness and efficacy of the ENN method, we intentionally enhanced statistical noise in the training data by the MD simulations, in which we terminated equilibration runs for all salt concentrations at 1 ns and began sampling. The dielectric constants calculated from the average of the total dipole moment over 2100 configurations serves as ground truth to examine the efficacy of our ensemble networks with BN and bagging, particularly when we substantially reduce the number of training data.

6.2.2 Design of sub-NN

For prediction of NaCl-containing water, we construct the ENN consisting of 5 identical sub-NNs with the topology of 2-50-50-50-50-1 (2 inputs, 4 hidden layers with

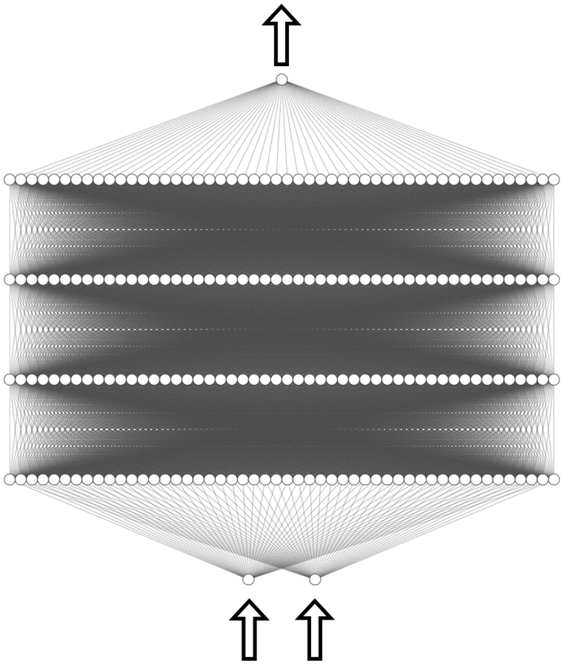


Figure 6.1: The sub-NN structure for NaCl-containing water. The inputs are the salt concentration c_s and the LJ parameter ϵ_{LJ} for the ion-ion interaction.

50 neurons per hidden layer, and 1 output) (see Fig. 6.1). Since the training data are more random, the number of neurons is increased to pass more information to the next layer, so that we constructed a shallower sub-NN with more neurons in hidden layers. The sequence of the activation function is (BN-ReLU)-Gauss-(BN-ReLU)-(BN-ReLU)-ReLU. To avoid the local minimum of training and accelerate the training speed, we use BN before each ReLU activation function. We assume that 40 samples are a reasonable number to maintain the Gaussian assumption in the BN. We also evaluate the model with the MSE loss function and perform the early stopping of training with a patience setting of 30 epochs. From the calculation of the MSE, we also empirically found that the sub-NN with 50 neurons per hidden layer serves

as a good architecture.

Hyper-parameter	Value or function
Number of subNN in an ENN	5
Number of hidden layers in a subNN	4
Number of activated neurons each hidden layer	50
Activation function in 1st hidden layer	Batch Normalization + ReLU
Activation function in 2nd hidden layer	Gaussian
Activation function in 3rd hidden layer	Batch Normalization + ReLU
Activation function in 4th hidden layer	Batch Normalization + ReLU
Activation function in output layer	ReLU
Patience of early-stopping	[15; 20; 25]
Batch size	[20; 30]
Train-test split ratio	4:1
Optimization algorithm	Adam optimizer
Learning rate in optimization algorithm	0.001
Cost or loss function each subNN use	Mean square error (Eq. 2.7)
Accuracy metrics	Mean absolute percentage error

Table 6.2

Hyper-parameters and their values or functions in the ENN model.

To increase the diversity of ENN, We combine sub-NNs with different patience and batch sizes. Details are as follows: sub-NN1 (patience = 15, batch size = 20), sub-NN2 (patience = 20, batch size=30), sub-NN3 (patience=20, batch size=20), sub-NN4 (patience=25, batch size=30), and sub-NN5 (patience=25, batch size=20). Hyper-parameters are summarized in Table 6.2.

#ENNs (ID number)	Training Error	Validation Error
1	**1015.5 (85.2%)**	**1164.9 (102.3%)**
2	1040.4 (85.4%)	1078.0 (104.0%)
3	1039.8 (85.1%)	1088.7 (102.6%)

Table 6.3
ENN training errors and validation errors. The topology of sub-NNs are all 2-50-50-50-50-1. The MAPE is shown in parentheses.

Sub-NN (No.1 ENN)	Training Error	Validation Error
1	1054.9 (90.1%)	1215.5 (92.7%)
2	1210.1 (103.1%)	1418.3 (107.3%)
3	1347.8 (97.2%)	1565.5 (99.86%)
4	1206.4 (96.0%)	1384.9 (87.9%)
5	1890.1 (101.6%)	2170.8 (105.5%)

Table 6.4
Sub-NNs training error and validation error of sub-NNs of No.1 ENN (bold one in Table 6.3). The topology of sub-NNs are all 2-50-50-50-50-1. The MAPE is shown in parentheses.

6.3 Results and discussion

Table 6.3 and Table 6.4 show the training and validation errors of our ENNs and sub-NNs, respectively. We also provide MAPE in parentheses. Note that unlike the case of the salt-free solvents, the training data consists of the values before taking the average and thus includes large statistical fluctuations. Accordingly, the training and validation errors appear to be large (Table 6.3 and Table 6.4).

The performance of ENNs consisting of five of these sub-NNs is shown in Table 5.3, we listed some other ENN examples, among them all sub-NNs' details of the bold row

(ENN1) are shown in Table 6.4. Here, ENN1 outperforms any of its sub-NNs. Unlike relatively large variations in the errors of sub-NNs in Table 6.4, the errors of ENNs are more stable. Thus, the training of ENNs is more robust than that of sub-NNs.

Because we use batch normalization in sub-NN, none of sub-NN is trapped in local minimum, so that each sub-NN are well-trained, which helps ENN to predict. Batch normalization increases the chance of jumping out of local minimum, and increasing the number sub-NNs does not significantly improve the ENN prediction.

The reason we do not use BN in previous chapter for salt-free solvents training is the limited number of training data. We calculate standard deviation while using BN, but the training data are statistical average values, so that the training data are not noisy. In that case, the BN cannot significantly enhance the prediction, and it also cannot significantly accelerate the speed of training. However, utilizing BN on 40 samples with noisy information can significantly accelerate the speed of training and obtain a good prediction by smoothing the noise.

Fig. 6.2 shows the ground truth and the prediction of ENN1, in which the salt concentration c_s and the LJ parameters ϵ_{LJ} for the ion-ion and ion-water interactions are changed. Note that training data at given salt concentration and LJ parameter (i.e., a data point) consists of only 40 configurations, although we need more training data to produce the ground truth. Thus, the training data may be disparate from the ground truth, yet the prediction compares favorably with the ground truth, i.e., the dielectric constant calculated from the average of the total dipole moment over 2100 configurations at each data point.

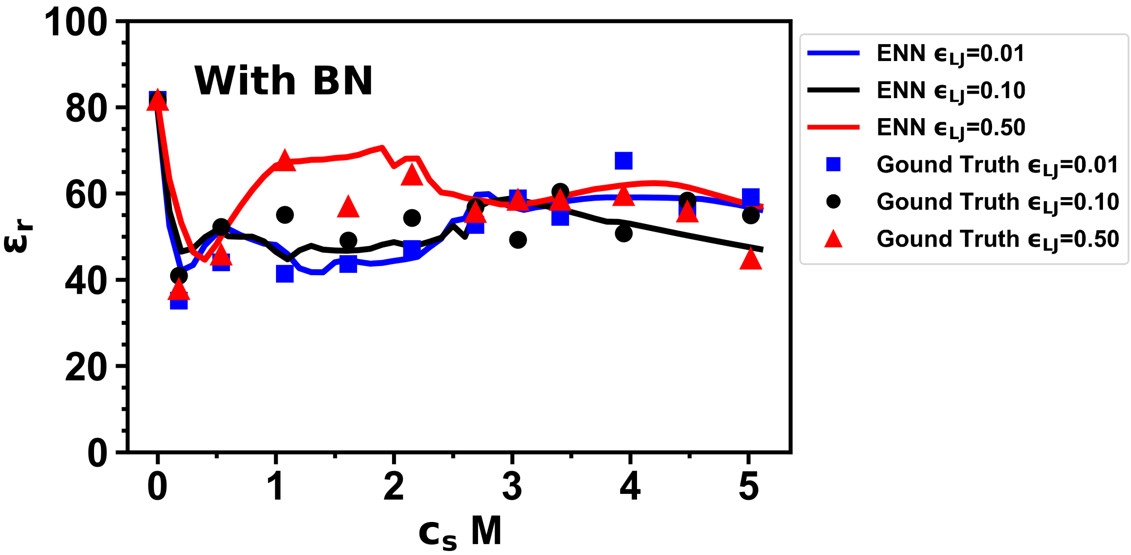


Figure 6.2: Comparison between the ground truth and the ENN prediction for the dielectric constant. Only BN was used. The solid lines indicate the predictions from the ENN for different salt concentrations c_s and different LJ parameters between the ions. “LJ” designates $\epsilon_{LJ} ion, ion$, where $\epsilon_{LJ}^{(ion,ion)}=\epsilon_{LJ}^{(+,+)}=\epsilon_{LJ}^{(+,-)}=\epsilon_{LJ}^{(-,-)}=\epsilon_{LJ}^{(water,+)}=\epsilon_{LJ}^{(water,-)}$. The symbols indicate the ground truth (the average over 2100 configurations at each data point). Reproduced from Ref. [2] with permission from the Materials Research Society.

6.3.1 Effect of batch normalization on prediction

Table 6.5 shows the difference in the dielectric constant between the prediction derived from the poor training data sets and the average calculated from the poor training data sets or the ground truth. Note that the predictions with BN are worse than those without BN when compared with the average calculated from the poor training

ϵ_{LJ} kcal/mol		40 configurations	2100 configurations
0.01	No BN	78.2 (15.7%)	65.1 (13.0%)
	BN	99.0 (16.4%)	**25.4 (8.8%)**
0.10	No BN	88.7 (15.1%)	130.1 (18.8%)
	BN	186.1 (18.6%)	**37.0 (9.8%)**
0.50	No BN	28.4 (11.1%)	79.8 (19.9%)
	BN	155.9 (20.9%)	**54.5 (10.9%)**

Table 6.5

MSEs (MAPEs in parentheses) of the prediction of the ENNs for NaCl-containing water with and without BN, relative to the average of the dielectric constant calculated from the poor training data sets and the ground truth. These results were calculated from Fig. 6.2 and Fig. 6.3.

data sets. In contrast, the performance of BN becomes better when compared with the ground truth, as highlighted in bold. This difference indicates that BN drives the prediction derived from the poor training data sets to approach the ground truth. Thus, BN qualitatively improves on training, tracking down the optimal values of the dielectric constant out of the noisy datasets.

To illustrate the role of BN, we show Fig. 6.3, in which the dielectric constants calculated from individual configurations widely spread around the mean values. Note that the averages calculated from the poor training data sets may not be close to the ground truth, thus they can be highly inaccurate values. Nevertheless, when we use BN, the predictions derived from the poor training data sets do not trace the false averages and instead tend to approach the ground truth. On the other hand, when we do not use BN, the predictions tend to approach the false averages. For example, this fact would be straightforwardly noticed at the data point for $\epsilon_{LJ} = 0.1$ kcal/mol and $c_s = 3.407$ M and at nearly all data points for $\epsilon_{LJ} = 0.5$ kcal/mol. Thus, BN-combined

ENNs appear to be robust against the significant reduction in the number of training data, and therefore, would considerably speed up the parameter adjustment and the exploration of new features in noisy data over vast parameter space.

6.3.2 Effect of bootstrap aggregating (bagging) on prediction

Finally, we also employed bagging, a treatment for decreasing variance to avoid overfitting. By randomly grouping training data, we obtained 5 different training datasets and then fed these training data into 5 different sub-NNs. Our results showed that bagging can outperform BN, but a hybrid of BN and bagging makes the prediction less accurate, as highlighted by bold figures in Table 6.6. In Fig. 6.4, the results obtained from the bagging-only method can more accurately capture the nonmonotonicity in low salt concentrations with all LJ parameters and the nonmonotonicity in high concentrations in the case of $\epsilon_J = 0.5$ kcal/mol. Nevertheless, the optimization of the bagging size (the number of training data in bags selected from original training data to create diverse samples) is crucial to outperforming BN. However, this optimization may not always be guaranteed or may require further computational expense because the training performance depends on various hyperparameters. Therefore, the addition of BN between the activation layers would be relatively straightforward to substantially improve the ENN performance.

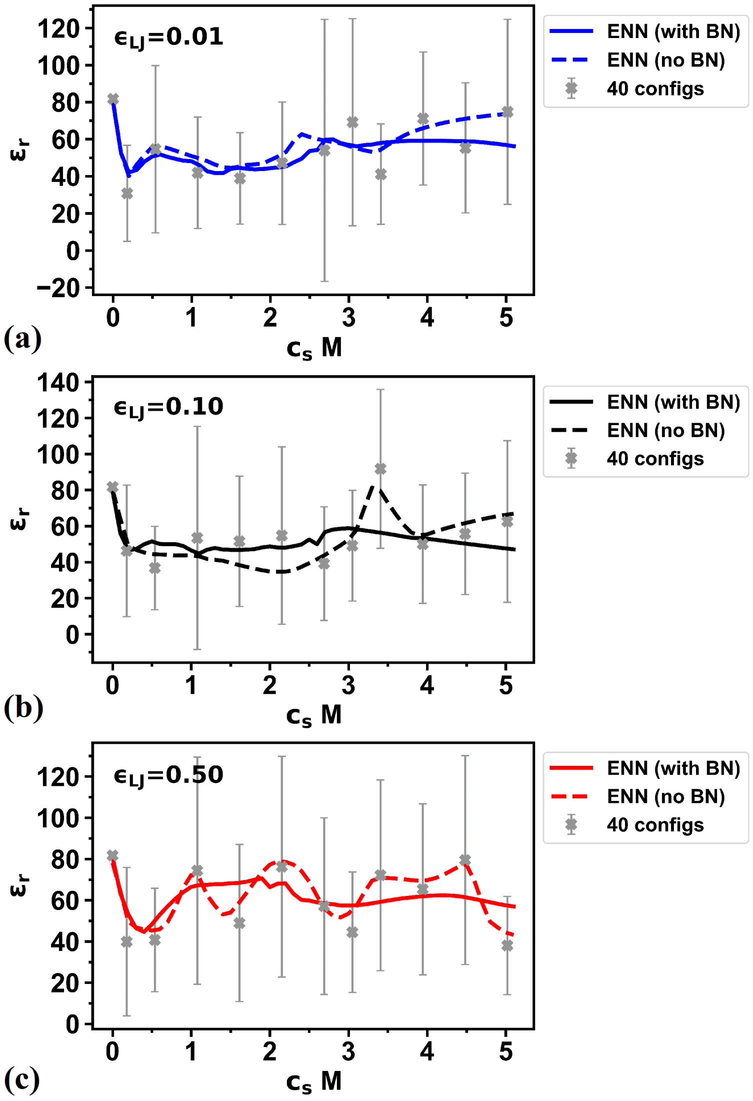


Figure 6.3: Efficacy of BN with (a) $\epsilon_{LJ} = 0.01$ kcal/mol, (b) $\epsilon_{LJ} = 0.10$ kcal/mol, and (c) $\epsilon_{LJ} = 0.50$ kcal/mol. The solid and dashed lines indicate the ENN predictions with BN and without BN, respectively. The grey symbols indicate the average over the 40 configurations at each data point, whereas the standard deviations are represented by error bars. The ENNs with BN improves on training. Reproduced from Ref. [2] with permission from the Materials Research Society.

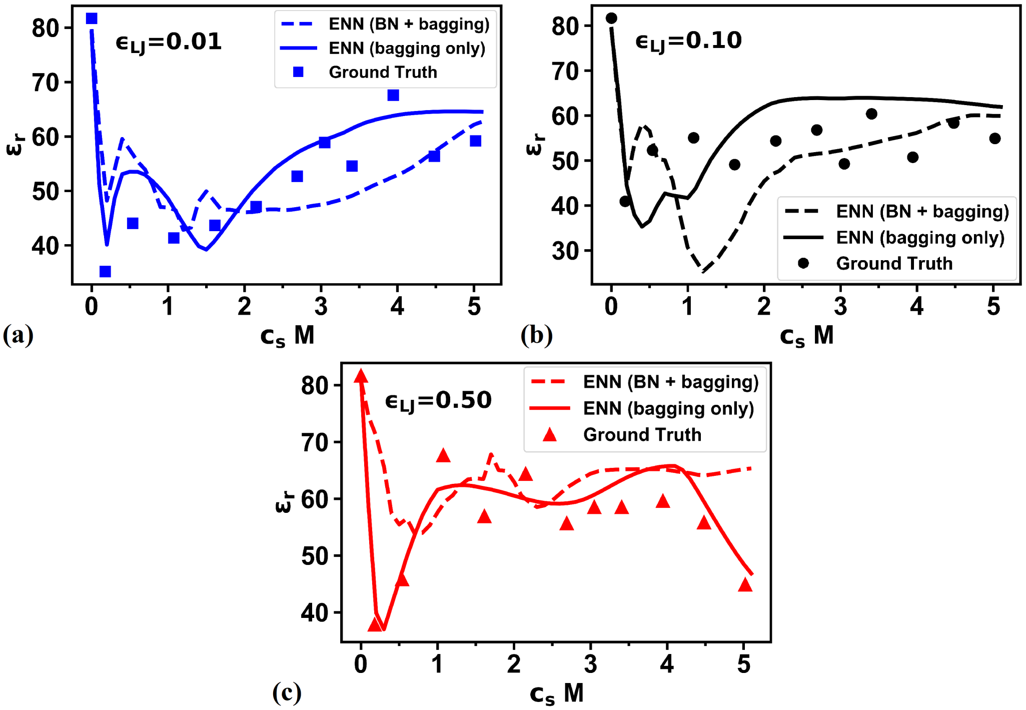
ϵ_{LJ} kcal/mol	Bagging	40 configurations	Ground truth
0.01	No BN	78.4 (14.7%)	31.1 (10.2%)
	BN	**132.5 (18.1%)**	**73.2 (14.1%)**
0.10	No BN	171.8 (18.3%)	88.6 (16.4%)
	BN	**245.3 (20.9%)**	97.0 (13.6%)
0.50	No BN	123.8 (16.1%)	17.2 (6.9%)
	BN	**285.7 (28.7%)**	**169.1 (20.5%)**

Table 6.6

MSEs (MAPEs) of the prediction of the ENNs for NaCl-containing water with and without BN under bagging, relative to the dielectric constant calculated from the poor training data sets and the ground truth. These results were calculated from Fig. 6.4.

6.3.3 Effect of combination method

Fig. 6.5 represents the results of a hybrid of five sub-NNs via a combination neural network (com-NN). Outputs from the five sub-NNs become inputs for the com-NN with three hidden layers consisting of 16 neurons each. All activation functions in the hidden layer are ReLUs. In this case, the com-NN worsens when the number of the neurons in the hidden layer increases to 30 or decreases to 10. Decreases in the number of the hidden layer did not provide any notable changes in the results. The results of the com-NN resemble those derived from the linear regression method. Thus, we suggest that the complexity of the hybridization for the output layer is unlikely to substantially improve the prediction.



(a) (b) (c)

Figure 6.4: BN + bagging (solid line) and bagging only (dashed line) with (a) $\epsilon_{LJ} = 0.01$ kcal/mol, (b) $\epsilon_{LJ} = 0.10$ kcal/mol, and (c) $\epsilon_{LJ} = 0.50$ kcal/mol. The symbols indicate the ground truth. The performance of the bagging-only method is tantamount to or somewhat exceeds that of BN-only method. Reproduced from Ref. [2] with permission from the Materials Research Society.

6.4 Conclusion

For salt-doped liquids, we trained ENNs with poor training data sets that are insufficient to produce the ground truth. Moreover, the training data widely spreads around the mean values at each data point. Moreover, the training data were widely distributed around the mean values for each MD parameter set. In this case, with

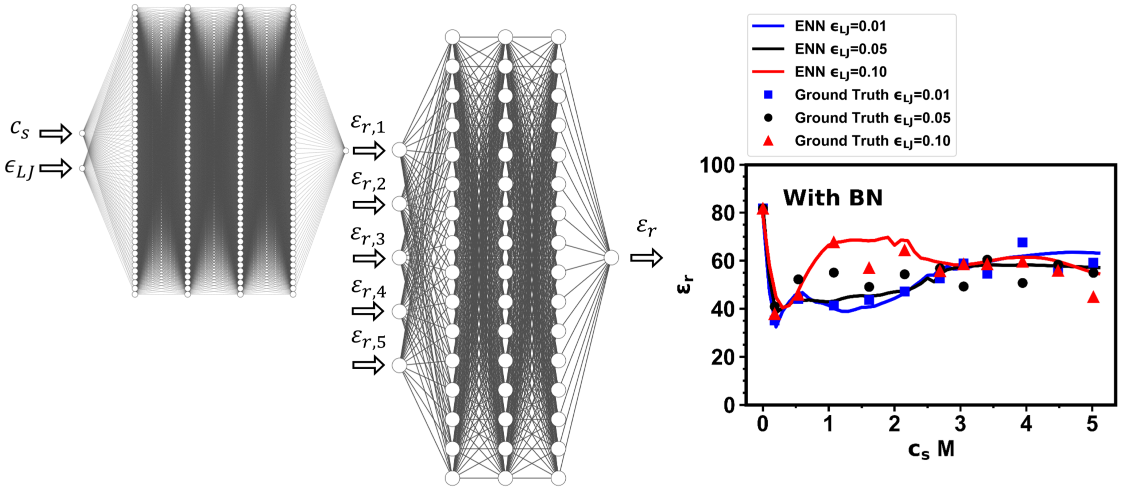


Figure 6.5: Dielectric constants of salt-doped solvents predicted by ENN1 consisting of sub-NNs combined through a neural network. Reproduced from Ref. [2] with permission from the Materials Research Society.

highly noisy data, BN or bagging enabled the predictions of ENNs to compare favorably with the full-simulation averages when employed independently. In other words, BN and bagging make the prediction of ENNs trace the optimal values derived from noisy datasets. Therefore, the ENNs are trained using 50-times less data than is used to calculate the dielectric constants from MD simulations. The predictions of ENNs trained on this small amount of data and using BN or bagging are in relatively good agreement with the full MD results.

Thus, we suggest that ENNs with BN or bagging substantially reduce the number of simulation runs and thus save computational effort. For example, the ENNs trained with a small number of possible input parameters can be used to choose values of the model parameters that best match target experiments, without having to run more MD simulations. The present methods would also be useful to simultaneously determine a number of model parameters for multicomponent systems.

Chapter 7

Unsupervised Learning on Cellulose Solubility in Ionic Liquids

7.1 Introduction

Unsupervised learning is discovering input-output relationships for unlabeled data. Several clustering and classification methods were compared including K-means clustering and hierarchical clustering on cellulose-containing systems. Talik et al. use machine learning tools for an empirical data analysis of the mutual interactions of water with cellulose polymers [105]. Random Forest from Machine learning algorithms is used to predict cellulose-rich material properties with ILs pretreatment [106]. Aromatic anions, such as salicylate anion, are known to improve the solubility of various types of cellulose [107, 108]. Nevertheless, interesting ILs have not been analyzed.

In this chapter, we will classify solubility phases in ion-containing liquids based 18 experimental data points. As we know, data from experiments are often significantly limited, because experiments are time-consuming and costly. Hence, extracting useful features from limited experimental data may be challenging. Since the data points are unlabeled for classification, unsupervised learning will provide an automatic and easy way to achieve this task.

7.2 Computational methods

I will combine supervised learning and unsupervised learning on this problem. The ENN discussed in the previous section is a supervised-learning algorithm. Then I'll apply the unsupervised learning, K-means clustering and agglomerative clustering for classifying the boundaries for phases.

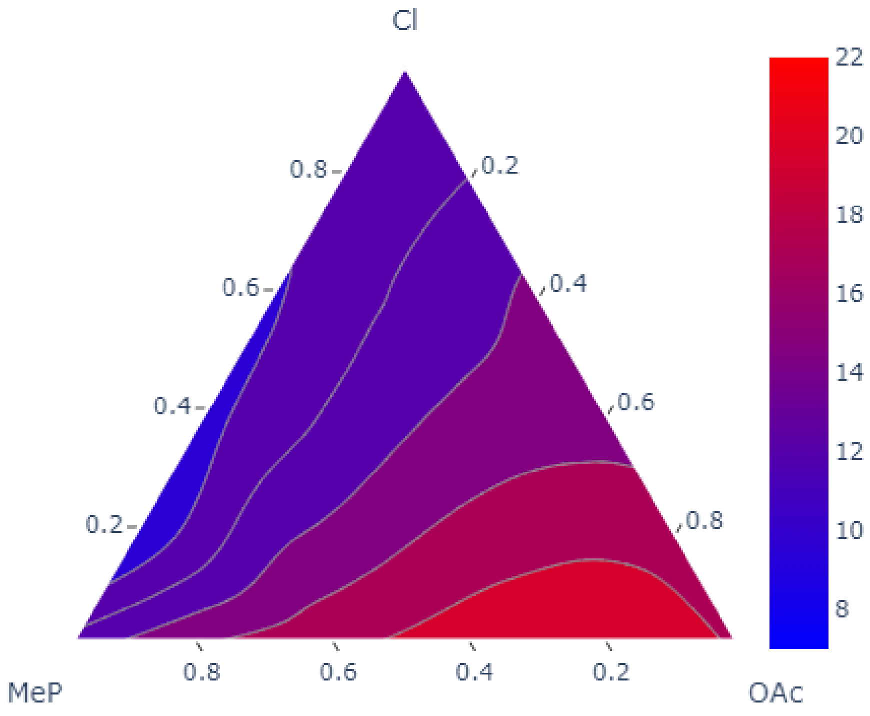


Figure 7.1: Ternary diagram of cellulose solubility based on experimental data. The color bar indicates cellulose solubility (wt%).

7.2.1 Data

Data is collected from our collaborator who works on solubility of cellulose in ILs mixtures. The ILs mixtures include one type of cation EMIM$^+$ and three types of anions (Cl$^-$, MeP$^-$, OAc$^-$). The features of data include molar ratio of three types of anions ([Cl], [MeP], [OAc]), and the cellulose solubility (wt%). Fig. 7.1 shows a ternary diagram for cellulose solubility changing with molar ratio of ions based on experimental data. The values of ion molar ratio are between 0.0 - 1.0. The range of solubility is 0 - 20 % according to the training data. The number of experimental data is 18. And all experimental data reside on the three axises. And the color within the triangle in Fig. 7.1 is interpolated, which is linearly estimated.

7.2.2 Design of ENN

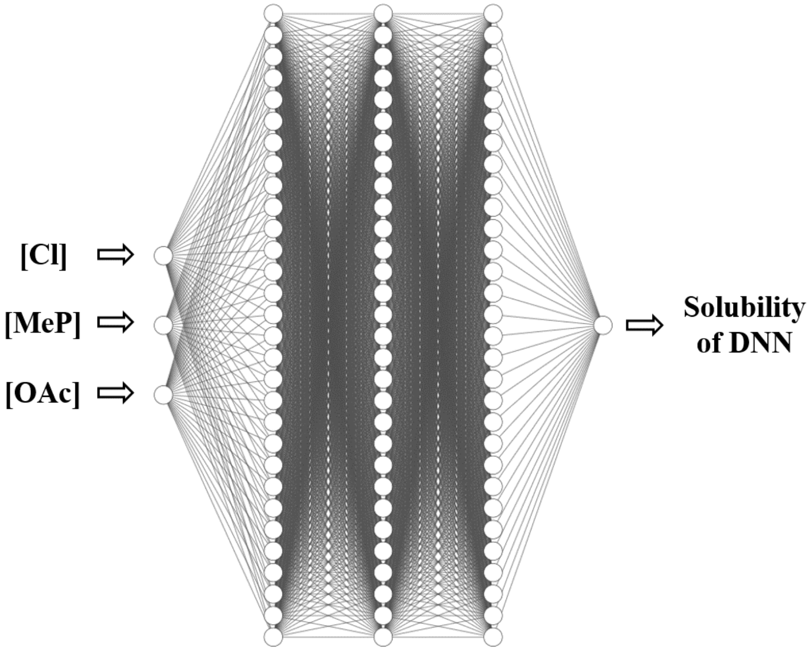


Figure 7.2: DNN structure for predicting cellulose solubility. This DNN has five input neurons and one output neuron. The DNN1, DNN2 have the identical structure with three input neurons ([Cl], [MeP] and [OAc]).

First, the deep neural network (DNN) I designed with three inputs ([Cl], [MeP], and [OAc]), 3 hidden layers with 30 neurons each layer, and one output. The activation function for all layers are ReLU. Through training, I found that the prediction of single DNN is not stable, the red region (low [Cl]) in Fig. 7.1 cannot be well predicted.

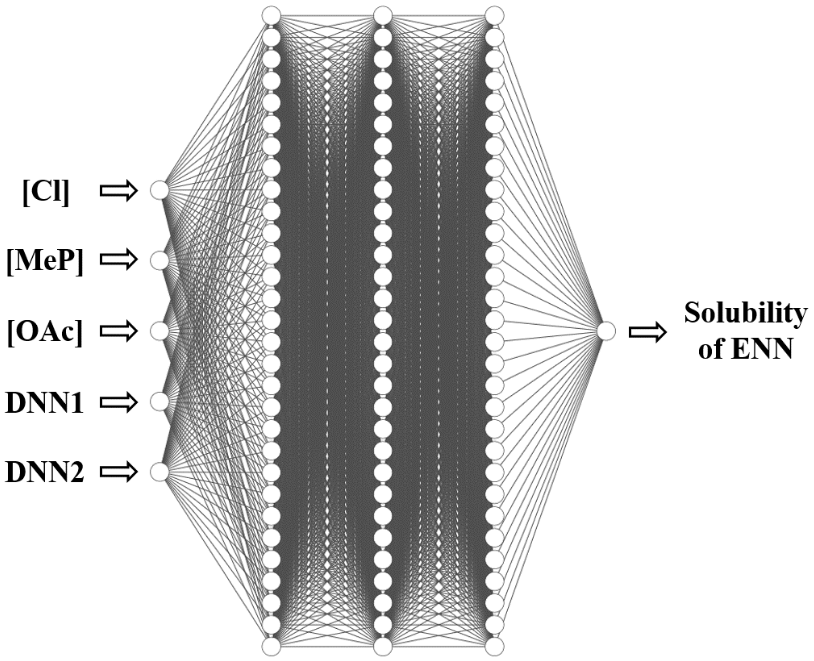


Figure 7.3: ENN structure for predicting cellulose solubility.This NN has five input neurons and one output neuron. The DNN1, DNN2 have the identical structure with ENN but only three input neurons ([Cl], [MeP] and [OAc]).

Then, I decide to construct an ENN to predict the phase behaviors of cellulose and ILs mixtures. It has been proven in previous chapters that an ENN outperforms a single NN. The structure of the ENN is shown in Fig. 7.2. The architecture of my ENN is 5-30-30-30-1, all activation function are ReLU. Three of inputs are molar ratio of three types of anions, the other two of inputs come from predictions of DNNs (Fig. 7.1). Both DNNs has identical structures. The architectures of ENN is 5-30-30-30-1. In addition, I also set up patience = 10 for early stopping to prevent over-fitting.

Hyper-parameter	Value or function
Number of input neurons in DNN	3
Number of input neurons in ENN	5
Number of hidden layers	3
Number of activated neurons each hidden layer	30
Activation function in hidden layers	ReLU
Activation function in output layer	ReLU
Patience of early-stopping	10
Batch size	6
Optimization algorithm	Adam optimizer
Learning rate in optimization algorithm	0.001
Cost or loss function each subNN use	Mean square error (Eq. 2.7)
Accuracy metrics	Mean absolute percentage error

Table 7.1

Hyper-parameters and their values or functions in the ENN model.

7.3 Results and discussions

Fig. 7.4 shows the prediction from ENN, we can see that the distribution of the cellulose solubility is very close to that of experimental data shown in Fig. 7.1. The color bar represents the cellulose solubility (wt%), red color indicates higher solubility, which happens in area around ([Cl], [MeP], [OAc]) = (0.0, 0.2, 0.8). Blue color indicates lower solubility, which happens in area around ([Cl], [MeP], [OAc]) = (0.3, 0.7, 0.0). Based on this ternary diagram, we believe that the ion [OAc] favors the cellulose solubility to some degree.

Fig. 7.5 shows results predicted by K-means clustering. K = 2 in Fig. 7.5(a), which means two clusters. One is the high solubility area, the other is the low solubility area.

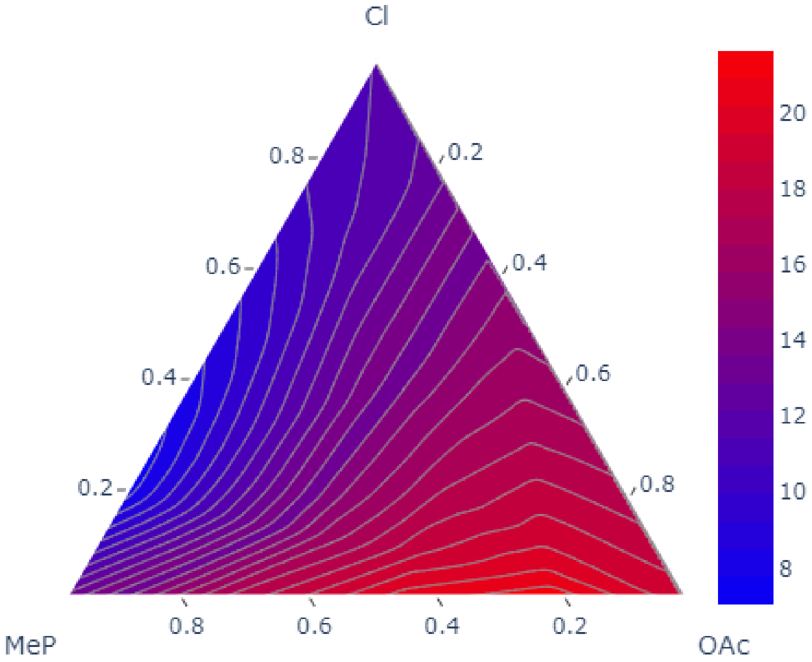
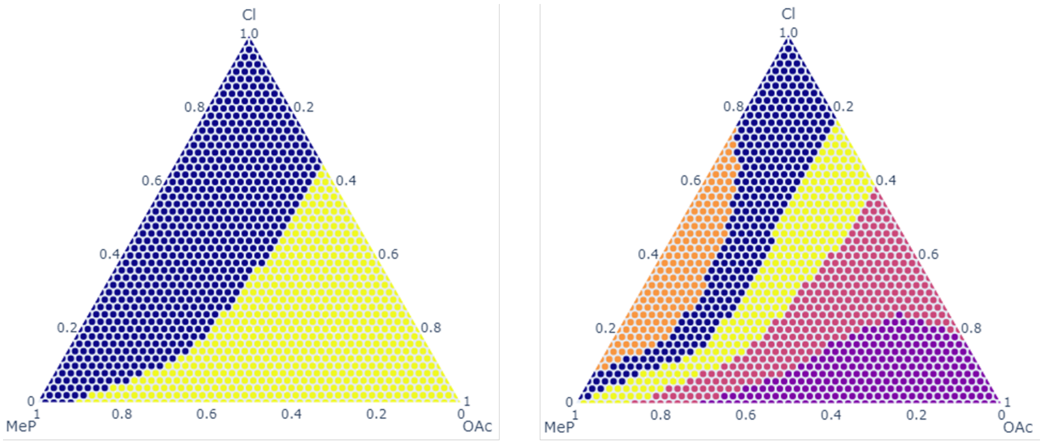


Figure 7.4: Ternary diagram of cellulose solubility based on ENN predictions. The color bar indicates cellulose solubility (wt%).

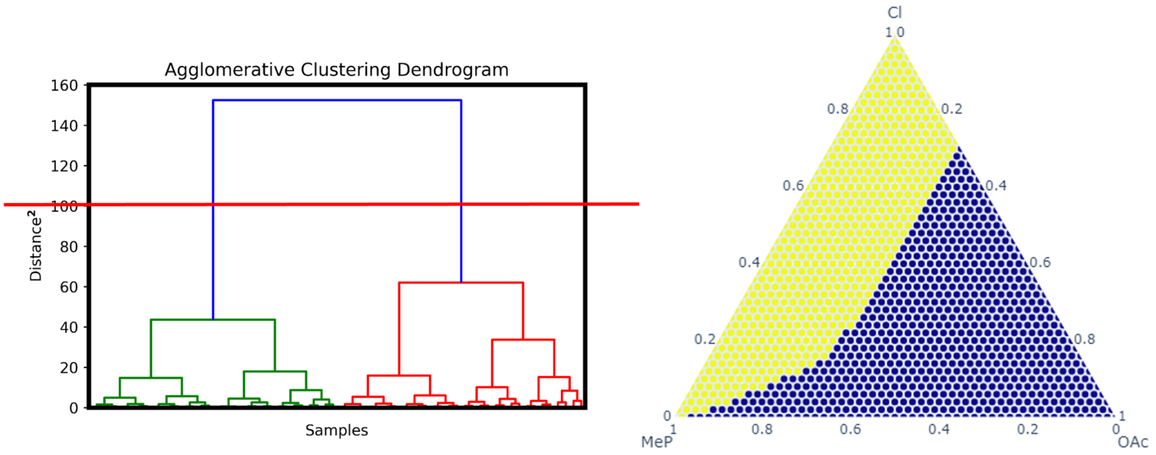
And in Fig. 7.5, K=5 that means five clusters. We can see five phases with different cellulose solubility in ILs. From orange at the left to purple at the right bottom corner, phases gradually change from the solubility-unfavored to solubility-favored.

In Fig. 7.6(b), results predicted by agglomerative clustering is quite similar with K-means clustering. The colors only represent clusters, because the color is arranged randomly. Two colors indicates two phases, even though the color is reversed between Fig. 7.5(a) and Fig. 7.6(b). I fixed the squared distance between clusters' centroids as



(a) (b)

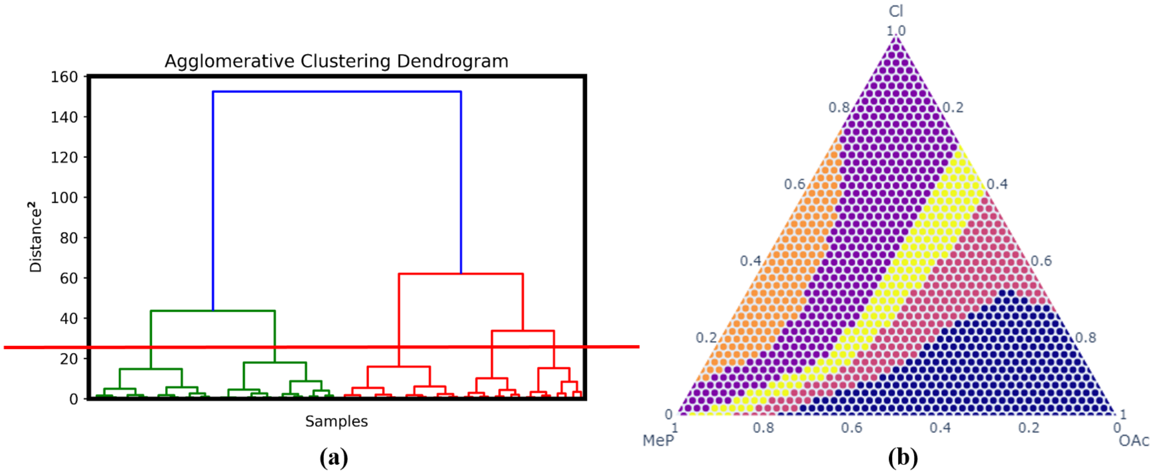
Figure 7.5: The boundary for solubility predicted by K-means clustering, where (a) K = 2 and (b) K = 5.



(a) (b)

Figure 7.6: (a) A dendrogram for agglomerative clustering. The red line is where to cluster, the number of interceptions indicates the number of clusters. In the diagram, 2 interception represents 2 clusters. (b) The boundary for solubility predicted by agglomerative clustering, which has 2 clusters.

the red line trims in Fig. 7.6(a), where the distance$^2 = 100$, the result turns out to be 2 classes. Within 10% solubility difference, samples are merged into one phase. If I move the red line toward to bottom in the dendrogram, where the distance$^2 = 25$ which



(a) (b)

Figure 7.7: (a) A dendrogram for agglomerative clustering. The red line is where to cluster, the number of interceptions indicates the number of clusters. In the diagram, 5 interception represents 5 clusters. (b) The boundary for solubility predicted by agglomerative clustering, which has 5 clusters.

means the distance between clusters' centroids become closer. Within 5% solubility difference, samples are merged into one phase. In Fig. 7.7(a), we can visually see more interceptions between the red line the tree-like hierarchy dendrogram, that is, the result turns out to have more clusters (Fig. 7.7(b)). And more clusters may imply more phases that not explored yet in this ternary solubility diagram.

7.4 Conclusion

In conclusion, by combining the supervised learning and the unsupervised learning, my machine learning algorithms are able to learn the ternary diagram and how many phases it is supposed to have from limited experimental data. Based on the ENN predictions, I can utilize clustering algorithms to decide the number of phases and to clarify their boundaries. We found strong evidence of a solubility-favored phase (left part of 2 clusters) and a solubility-unfavored phase (right part of 2 clusters), and know the relation between the ion concentration and the cellulose solubility. Cellulose solubility increases as [Cl] decreases, [OAc] promote solubility, and [MeP] does not have significant impact on solubility.

In addition, the number of phases has to be specified when using K-means clustering, while the number of phases is not required for agglomerative clustering. And the phases predicted by K-means clustering may be different on different runs because of the random initialization of points. However, agglomerative clustering always predict the same phases.

Chapter 8

Summary and Future Perspectives

8.1 Summary

We have developed the DLA model proposed by Niemeyer [26] and later developed for zinc dendrite growth by Chen and Jorné [24]. The present study was also motivated by our recent experimental observation regarding remarkably uniform, unconventional lithium dendrite growth in Fig. 3.11. Despite the simple algorithm of the growth patterns, the original toy model accounts for the fractal dimensions observed for dielectric breakdown and zinc dendrite. In other words, the model contains an adjustable parameter that controls the growth probability, but it can be fixed when the fractal dimension is determined in experiments. The present study relies on the strong applicability of the model to the pattern formation due to the electrostatic fields. Thus, we applied the algorithm of the pattern formation to metal dendrite growth in salt-containing liquids between the conducting plates. The main conclusion in our study is summarized as follows: (1) The addition of salt ions can significantly flatten the dendrite shape and inhibit the dendrite growth, primarily due to electrostatic shielding near the pointy regions of the dendrite. (2) It is critical to weaken ion pairing and clustering in electrolytes to inhibit the dendrite growth. Our results show that small salt ions appear to have difficulty achieving this requirement and lead to the undesirable non-monotonic behavior of the dendrite growth as a function of the salt concentration. However, large salts such as ILs can be dissociated near the operating temperatures to a great degree to solve this issue. (3) The size asymmetry that occurs when either the cation or the anion is large (4 Å or larger) affects the dendrite's shape and growth, primarily because ion pairing and clustering are significantly inhibited. In this case, the inhibition of the dendrite growth is not significantly affected by changes in the dielectric constant and applied voltage. The overall conclusion from (1) to (3) remains unchanged when $0 \leq \eta \leq 2$, and appears to hold true in general. Thus, we suggest that ILs be a prospective material to inhibit metal dendrite formation. Among others, the present simulations suggest that ILs with large size asymmetry between the cation and anion serve as a good electrolyte in lithium-ion batteries, as consistent with our experimental observation for lithium deposits under 1M $LiPF_6$ in EC/DMC + 10% [BMIM][TFSI] in experiments. Concomitantly, the 2D growth appears to be more constrained than the 3D growth. In the case of a 3D model, the dendrite growth in the new lateral direction may easily occur when the ions inhibit the longitudinal growth of the dendrite toward the cathode. Still, this lateral growth appears to facilitate the uniformity of the dendrite surface. Thus, we anticipate that our overall conclusion remains unchanged even in a 3D model. Finally, we considered a surrogate model that accounts for the MC simulations using ENNs. In our study, we found that a single NN can be consistent with the MC simulation results, but ENNs with sub-NNs can be more robust against the choice of the activation functions and can easily be trained from various initial weights. This is mainly because the ENNs can be trained well when poor sub-NNs are combined with good sub-NNs. Importantly, unlike 200 samples for the average of the MC simulation results for each volume fraction (i.e., each data point), we used only 20 samples for training. Thus, our surrogate model serves as an alternative to the computationally demanding MC simulations.

Subsequently, we constructed ENNs in machine-learning methods for the calculation of the dielectric constant of salt-free and salt-doped liquids. For training data, we used SF MD simulation methods that account for the dielectric response of dipolar fluids. We showed that our ENNs with the activation layers containing sine and gaussian functions capture the trends of the dielectric constant of seven salt-free solvents. For salt-doped liquids, we trained ENNs with poor training data sets that are not enough to produce the ground truth. Moreover, the training data widely spreads around the mean values at each data point. In this case with highly noisy data, BN or bagging substantially enables the prediction of ENNs to compare favorably with the ground truth when employed independently. In other words, BN and bagging make the prediction of ENNs trace the optimal values out of noisy datasets. Thus, we suggest that ENNs with BN or bagging substantially reduce the number of simulation measurements and thus save computational expense. The present method would also be useful to simultaneously determine a number of model parameters for a multicomponent system.

Then we combined unsupervised learning with supervised ensemble learning for predicting the ternary behaviors of cellulose and ILs mixtures. With hierarchical clustering, the number of phases is not necessary to be specified, and the prediction by hierarchical clustering is more stable than K-means clustering because it does not rely on random initialization.

8.2 Future perspectives

There are challenges for machine learning algorithms but their usefulness, accuracy and their ability to take in the battery characteristics as inputs cannot be ignored. This work can be extended to other battery chemistry. Only four inputs are used in this study, more battery characteristics can be incorporated as inputs for better lithium dendrite suppression. And surrogate model architecture designs will give an idea to MC simulations, MD simulations, and phase-field models that have large number of tunable parameters, to accelerate simulation process with higher accuracy. Artificial neural networks provide prediction values. Additionally, we can use Gaussian process or Bayesian neural networks to provide predictions with uncertainty. However, implementing a Bayesian neural network requires more statistical and mathematical expertise.

Before the machine learning era, research on properties of soft-matter systems has employed many of the well-established physical and mathematical ideas, even though now machine learning is also developed based on mathematical algorithms. The fitting of model parameters sometimes has been seen more inconvenient than the crucial model generation process in research, and the machine learning helps reduce the workload of parameter fitting.

The research on physical, thermodynamic, and electrochemical properties of soft-matter systems could not only be based on numerical data, but also could be graph/image data. Deploying machine learning on these graph/image data could be efficient. Artificial neural networks are only one application of machine learning techniques, unsupervised learning, reinforcement learning, and other machine learning techniques can also provide new paths for soft-matter physics research. However, physical knowledge is the key to solving the issues. Meanwhile, the black-box nature of machine learning techniques will be greatly diminished along with the addition of physical knowledge.

References

[1] Gao, T.; Qian, Z.; Chen, H.; Shahbazian-Yassar, R.; Nakamura, I. *Molecular Systems Design & Engineering* **2022**, *7*(3), 260–272.

[2] Gao, T.; Shock, C. J.; Stevens, M. J.; Frischknecht, A. L.; Nakamura, I. *MRS Communications* **2022**, *12*, 966–974.

[3] Hagopian, A.; Doublet, M.-L.; Filhol, J.-S. *Energy & Environmental Science* **2020**, *13*(12), 5186–5197.

[4] Kim, W.-S.; Yoon, W.-Y. *Electrochimica Acta* **2004**, *50*(2), 541–545.

[5] Yamaki, J.-i.; Tobishima, S.-i.; Hayashi, K.; Keiichi, S.; Nemoto, Y.; Arakawa, M. *Journal of Power Sources* **1998**, *74*(2), 219–227.

[6] Xu, K. *Chemical Reviews* **2004**, *104*(10), 4303–4417.

[7] Liu, Y.; Liu, Q.; Xin, L.; Liu, Y.; Yang, F.; Stach, E. A.; Xie, J. *Nature Energy* **2017**, *2*(7), 17083.

[8] Rosso, M.; Brissot, C.; Teyssot, A.; Dollé, M.; Sannier, L.; Tarascon, J.-M.; Bouchet, R.; Lascaud, S. *Electrochimica Acta* **2006**, *51*(25), 5334–5340.

[9] Meng, X.; Xu, Y.; Cao, H.; Lin, X.; Ning, P.; Zhang, Y.; Garcia, Y. G.; Sun, Z. *Green Energy & Environment* **2020**, *5*(1), 22–36.

[10] Feynman, R.; Leighton, R.; Sands, M. *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Mainly Electromagnetism and Matter;* Basic Books, 2015.

[11] Haynes, W. M.; Lide, D. R.; Bruno, T. J. *CRC handbook of chemistry and physics;* CRC press, 2016.

[12] Chan, H.; Cherukara, M. J.; Narayanan, B.; Loeffler, T. D.; Benmore, C.; Gray, S. K.; Sankaranarayanan, S. K. R. S. *Nature Communications* **2019**, *10*(1), 379.

[13] Liu, Y.; Hong, W.; Cao, B. *Energy* **2019**, *188*, 116091.

[14] Deng, J.; Jia, G. *Fluid Phase Equilibria* **2022**, *561*, 113545.

[15] Hutchinson, S. T.; Kobayashi, R. *Journal of Chemical Information and Modeling* **2019**, *59*(4), 1338–1346.

[16] Shi, Y.; Lam, S. T.; Beck, T. *Chem. Sci.* **2022**, *13*, 8265–8273.

[17] Chen, Y.; Krämer, A.; Charron, N. E.; Husic, B. E.; Clementi, C.; Noé, F. *The Journal of Chemical Physics* **2021**, *155*(8), 084101.

[18] Wei, Y.; Chin, K.; Barge, L. M.; Perl, S.; Hermis, N.; Wei, T. *The Journal of Physical Chemistry B* **2020**, *124*(50), 11491–11500.

[19] Alshehri, M.; Kumar, M.; Bhardwaj, A.; Mishra, S.; Gyani, J. *Water* **2021**, *13*(9).

[20] Tovey, S.; Narayanan Krishnamoorthy, A.; Sivaraman, G.; Guo, J.; Benmore, C.; Heuer, A.; Holm, C. *The Journal of Physical Chemistry C* **2020**, *124*(47), 25760–25768.

[21] Sun, Y.; Chen, M.; Zhao, Y.; Zhu, Z.; Xing, H.; Zhang, P.; Zhang, X.; Ding, Y. *Journal of Molecular Liquids* **2021**, *333*, 115970.

[22] Dhakal, P.; Shah, J. K. *Fluid Phase Equilibria* **2021**, *549*, 113208.

[23] Chandrasekaran, A.; Kim, C.; Venkatram, S.; Ramprasad, R. *Macromolecules* **2020**, *53*(12), 4764–4769.

[24] Chen, C. P.; Jorne, J. *Journal of the Electrochemical Society* **1990**, *137*(7), 2047–2051.

[25] Aryanfar, A.; Brooks, D.; Merinov, B. V.; Goddard, W. A.; Colussi, A. J.; Hoffmann, M. R. *The Journal of Physical Chemistry Letters* **2014**, *5*(10), 1721–1726.

[26] Niemeyer, L.; Pietronero, L.; Wiesmann, H. J. *Physical Review Letters* **1984**, *52*(12), 1033–1036.

[27] Meakin, P. *Fractals, scaling and growth far from equilibrium*, Vol. 5; Cambridge university press, 1998.

[28] Dominkovics, C.; Harsányi, G. In *33rd International Spring Seminar on Electronics Technology*, pages 200–203, 2010.

[29] Ioffe, S.; Szegedy, C. *CoRR* **2015**, *abs/1502.03167*.

[30] keras. Chollet, F.; others. **2015**.

[31] Tang, Y.; Guo, Y.; Sun, Q.; Tang, S.; Li, J.; Guo, L.; Duan, J. *Optik* **2018**, *165*, 179–185.

[32] Kurita, H.; Suganuma, M.; Wang, Y.; Narita, F. *Advanced Engineering Materials* **2022**, *24*(5), 2101072.

[33] Shrifan, N. H. M. M.; Jawad, G. N.; Isa, N. A. M.; Akbar, M. F. *IEEE Access* **2021**, *9*, 4820–4828.

[34] Lukasiak, B. M.; Faria, R.; Zomer, S.; Brereton, R. G.; Duncan, J. C. *Analyst* **2006**, *131*, 73–80.

[35] Partheniadis, I.; Toskas, M.; Stavras, F.-M.; Menexes, G.; Nikolakakis, I. *Processes* **2020**, *8*(10).

[36] Armand, M.; Endres, F.; MacFarlane, D. R.; Ohno, H.; Scrosati, B. *Nature Materials* **2009**, *8*(8), 621–629.

[37] Maniam, K. K.; Paul, S. **2020**, *10*(15), 5321.

[38] Matsui, M. *Journal of Power Sources* **2011**, *196*(16), 7048–7055.

[39] Bucur, C. B.; Gregory, T.; Oliver, A. G.; Muldoon, J. *Journal of Physical Chemistry Letters* **2015**, *6*(18), 3578–3591.

[40] Song, J.; Lee, H.; Choo, M. J.; Park, J. K.; Kim, H. T. *Scientific Reports* **2015**, *5*.

[41] Selim, R.; Bro, P. *Journal of the Electrochemical Society* **1974**, *121*(11), 1457–1459.

[42] Parker, J. F.; Chervin, C. N.; Pala, I. R.; Machler, M.; Burz, M. F.; Long, J. W.; Rolison, D. R. *Science* **2017**, *356*(6336), 414–417.

[43] Howlett, P. C.; MacFarlane, D. R.; Hollenkamp, A. F. *Electrochemical and Solid-State Letters* **2004**, *7*(5), A97.

[44] Monroe, C.; Newman, J. *Journal of the Electrochemical Society* **2005**, *152*(2), A396–A404.

[45] Hallinan, D. T.; Mullin, S. A.; Stone, G. M.; Balsara, N. P. *Journal of the Electrochemical Society* **2013**, *160*(3), A464–A470.

[46] Choudhury, S.; Mangal, R.; Agrawal, A.; Archer, L. A. *Nature Communications* **2015**, *6*.

[47] Li, W. Y.; Yao, H. B.; Yan, K.; Zheng, G. Y.; Liang, Z.; Chiang, Y. M.; Cui, Y. *Nature Communications* **2015**, *6*.

[48] Schauser, N. S.; Harry, K. J.; Parkinson, D. Y.; Watanabe, H.; Balsara, N. P. *Journal of the Electrochemical Society* **2015**, *162*(3), A398–A405.

[49] Tan, J. W.; Tartakovsky, A. M.; Ferris, K.; Ryan, E. M. *Journal of the Electrochemical Society* **2016**, *163*(2), A318–A327.

[50] Arora, P.; Zhang, Z. M. *Chemical Reviews* **2004**, *104*(10), 4419–4462.

[51] Wu, H.; Zhuo, D.; Kong, D. S.; Cui, Y. *Nature Communications* **2014**, *5*.

[52] Monroe, C.; Newman, J. *Journal of the Electrochemical Society* **2003**, *150*(10), A1377–A1384.

[53] Mayers, M. Z.; Kaminski, J. W.; Miller, T. F. *Journal of Physical Chemistry C* **2012**, *116*(50), 26214–26221.

[54] Abbott, A. P.; McKenzie, K. J. *Physical Chemistry Chemical Physics* **2006**, *8*(37), 4265–4279.

[55] Basile, A.; Bhatt, A. I.; O'Mullane, A. P. *Nature Communications* **2016**, *7*.

[56] Pearson, A. K.; Kao, P.; O'Mullane, A. P.; Bhatt, A. I. *Physical Chemistry Chemical Physics* **2017**, *19*(22), 14745–14760.

[57] Momeni, K.; Ji, Y.; Wang, Y.; Paul, S.; Neshani, S.; Yilmaz, D. E.; Shin, Y. K.; Zhang, D.; Jiang, J.-W.; Park, H. S.; Sinnott, S.; van Duin, A.; Crespi, V.; Chen, L.-Q. *npj Computational Materials* **2020**, *6*(1), 22.

[58] Schulze, T. P. *Physical Review E* **2008**, *78*(2).

[59] Krumwiede, T. R.; Schulze, T. P. *Modelling and Simulation in Materials Science and Engineering* **2017**, *25*(2).

[60] Selis, L. A.; Seminario, J. M. *RSC Adv.* **2019**, *9*, 27835–27848.

[61] Wheeler, A. A.; Murray, B. T.; Schaefer, R. J. *Physica D* **1993**, *66*(1-2), 243–262.

[62] Karma, A.; Rappel, W. J. *Physical Review E* **1998**, *57*(4), 4323–4349.

[63] Boettinger, W. J.; Warren, J. A.; Beckermann, C.; Karma, A. *Annual Review of Materials Research* **2002**, *32*, 163–194.

[64] Chen, L. Q. *Annual Review of Materials Research* **2002**, *32*, 113–140.

[65] Shibuta, Y.; Okajima, Y.; Suzuki, T. *Science and Technology of Advanced Materials* **2007**, *8*(6), 511–518.

[66] Qin, R. S.; Bhadeshia, H. K. *Materials Science and Technology* **2010**, *26*(7), 803–811.

[67] Liang, L. Y.; Qi, Y.; Xue, F.; Bhattacharya, S.; Harris, S. J.; Chen, L. Q. *Physical Review E* **2012**, *86*(5).

[68] Ely, D. R.; Jana, A.; Garcia, R. E. *Journal of Power Sources* **2014**, *272*, 581–594.

[69] Liang, L. Y.; Chen, L. Q. *Applied Physics Letters* **2014**, *105*(26).

[70] Chen, L.; Zhang, H. W.; Liang, L. Y.; Liu, Z.; Qi, Y.; Lu, P.; Chen, J.; Chen, L. Q. *Journal of Power Sources* **2015**, *300*, 376–385.

[71] Cogswell, D. A. *Physical Review E* **2015**, *92*(1).

[72] Jana, A.; Ely, D. R.; García, R. E. *Journal of Power Sources* **2015**, *275*, 912–921.

[73] Wang, K. L.; Pei, P. C.; Ma, Z.; Chen, H. C.; Xu, H. C.; Chen, D. F.; Wang, X. Z. *Journal of Materials Chemistry A* **2015**, *3*(45), 22648–22655.

[74] Foroozan, T.; Soto, F. A.; Yurkiv, V.; Sharifi-Asl, S.; Deivanayagam, R.; Huang, Z.; Rojaee, R.; Mashayek, F.; Balbuena, P. B.; Shahbazian-Yassar, R. **2018**, *28*(15), 1705917.

[75] Yurkiv, V.; Foroozan, T.; Ramasubramanian, A.; Shahbazian-Yassar, R.; Mashayek, F. *MRS Communications* **2018**, *8*(3), 1285–1291.

[76] Yurkiv, V.; Foroozan, T.; Ramasubramanian, A.; Shahbazian-Yassar, R.; Mashayek, F. *Electrochimica Acta* **2018**, *265*, 609–619.

[77] Gould, H.; Tobochnik, J.; Meredith, D. C.; Koonin, S. E.; McKay, S. R.; Christian, W. *Computers in Physics* **1996**, *10*(4), 349–349.

[78] Wolfram, S. *Nature* **1984**, *311*(5985), 419–424.

[79] Kardar, M.; Parisi, G.; Zhang, Y.-C. *Physical Review Letters* **1986**, *56*(9), 889–892.

[80] Brú, A.; Pastor, J. M.; Fernaud, I.; Brú, I.; Melle, S.; Berenguer, C. *Physical Review Letters* **1998**, *81*(18), 4008–4011.

[81] Brú, A.; Albertos, S.; Luis Subiza, J.; García-Asenjo, J. L.; Brú, I. *Biophysical Journal* **2003**, *85*(5), 2948–2961.

[82] Eden, M. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, Vol. 4, pages 223–239. University of California Press Berkeley, 1961.

[83] Borukhov, I.; Andelman, D.; Orland, H. *Electrochimica Acta* **2000**, *46*(2), 221–229.

[84] Bazant, M. Z.; Storey, B. D.; Kornyshev, A. A. *Physical Review Letters* **2011**, *106*(4), 046102.

[85] Wu, B.; Han, S.; Shin, K. G.; Lu, W. *Journal of Power Sources* **2018**, *395*, 128–136.

[86] Quartulli, M.; Gil, A.; Florez-Tapia, A. M.; Cereijo, P.; Ayerbe, E.; Olaizola, I. G. *Energies* **2021**, *14*(14), 4115.

[87] Clemen, R. T. *International Journal of Forecasting* **1989**, *5*(4), 559–583.

[88] Wolpert, D. H. *Neural networks* **1992**, *5*(2), 241–259.

[89] Zhang, X.; Mesirov, J. P.; Waltz, D. L. *Journal of Molecular Biology* **1992**, *225*(4), 1049–1063.

[90] Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. *Annual Review of Physical Chemistry* **2020**, *71*(1), 361–390.

[91] Wang, J.; Olsson, S.; Wehmeyer, C.; Pérez, A.; Charron, N. E.; de Fabritiis, G.; Noé, F.; Clementi, C. *ACS Central Science* **2019**, *5*(5), 755–767.

[92] Krishnamoorthy, A.; Nomura, K.; Baradwaj, N.; Shimamura, K.; Rajak, P.; Mishra, A.; Fukushima, S.; Shimojo, F.; Kalia, R.; Nakano, A.; Vashishta, P. *Physical Review Letters* **2021**, *126*(21).

[93] Ghorbanfekr, H.; Behler, J.; Peeters, F. M. *Journal of Physical Chemistry Letters* **2020**, *11*(17), 7363–7370.

[94] Gu, Q. Q.; Zhang, L. F.; Feng, J. *Science Bulletin* **2022**, *67*(1), 29–37.

[95] Floyd, J. E.; Lukes, J. R. *Journal of Chemical Physics* **2022**, *156*(18).

[96] Reid, D. J. *Economica* **1968**, *35*(140), 431–444.

[97] Hansen, L. K.; Salamon, P. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1990**, *12*(10), 993–1001.

[98] Hashem, S.; Schmeiser, B.; Yih, Y. *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94) 3*, 1507–1512 vol.3.

[99] Le, H. M.; Raff, L. M. *Journal of Physical Chemistry A* **2010**, *114*(1), 45–53.

[100] Singh, S. K.; Bejagam, K. K.; An, Y. X.; Deshmukh, S. A. *Journal of Physical Chemistry A* **2019**, *123*(24), 5190–5198.

[101] Latino, D. A. R. S.; Freitas, F. F. M.; Aires-De-Sousa, J.; Fernandes, F. M. S. S. *International Journal of Quantum Chemistry* **2007**, *107*(11), 2120–2132.

[102] Shock, C. J.; Stevens, M. J.; Frischknecht, A. L.; Nakamura, I. *The Journal of Physical Chemistry B* **2020**, *124*(22), 4598–4604.

[103] Raju, G. G. *Dielectrics in electric fields: Tables, Atoms, and Molecules;* CRC press, 2017.

[104] Parascandolo, G.; Huttunen, H.; Virtanen, T. **2016**.

[105] Talik, P.; Mendyk, A. *Applied Sciences* **2021**, *11*(16).

[106] Phromphithak, S.; Onsree, T.; Tippayawong, N. *Bioresource Technology* **2021**, *323*, 124642.

[107] Brehm, M.; Pulst, M.; Kressler, J.; Sebastiani, D. *The Journal of Physical Chemistry B* **2019**, *123*(18), 3994–4003.

[108] Brehm, M.; Radicke, J.; Pulst, M.; Shaabani, F.; Sebastiani, D.; Kressler, J. *Molecules* **2020**, *25*(15).

[109] Selis, L. A.; Seminario, J. M. *Rsc Advances* **2019**, *9*(48), 27835–27848.

[110] Jung, H.; Yethiraj, A. *Journal of Physical Chemistry B* **2020**, *124*(41), 9230–9238.

[111] Nakamura, I. *New Journal of Physics* **2020**, *22*(1), 015001.

[112] Chen, M.; Li, J.; Tian, B.; Al-Hadeethi, Y. M.; Arkook, B.; Tian, X.; Zhang, X. **2021**, *9*(8), 87.

[113] Cubuk, E.; Schoenholz, S. S.; Rieser, J. M.; Malone, B. D.; Rottler, J.; Durian, D. J.; Kaxiras, E.; Liu, A. J. *Physical Review Letters* **2015**, *114*(10), 108001.

[114] Ferguson, A. L. *Journal of Physics: Condensed Matter* **2017**, *30*(4), 043002.

[115] Kadupitiya, J.; Fox, G. C.; Jadhao, V. Computational Science – ICCS, pages 116–130. Springer International Publishing, 2019.

[116] Liang, D.; Song, M.; Niu, Z.; Zhang, P.; Rafailovich, M.; Deng, Y. *MRS Advances* **2021**, *6*(13), 362–367.

[117] Tsai, S.-T.; Kuo, E.-J.; Tiwary, P. *Nature Communications* **2020**, *11*(1), 5115.

[118] Ulmer Ii, C. W.; Smith, D. A.; Sumpter, B. G.; Noid, D. I. *Computational and Theoretical Polymer Science* **1998**, *8*(3), 311–321.

[119] Maqsood, I.; Khan, R.; Abraham, A. *J Neural Comput. Appl.* **2004**, *13*(2), 112–122.

[120] Zhang, J.; Morris, A. J.; Martin, E. B.; Kiparissides, C. *Chemical Engineering Journal* **1998**, *69*(2), 135–143.

Appendix A

List of Selected Publications

1. **Gao, T.**, Shock, C. J., Stevens, M. J., Frischknecht, A. L., and Nakamura, I. (2022). Surrogate Molecular Dynamics Simulation Model for Dielectric Constants with Ensemble Neural Networks, *MRS Communications*, 12, 966–974.

2. **Gao, T.**, Qian, Z., Chen, H., Shahbazian-Yassar, R., and Nakamura, I. (2022). Inhibition of Lithium Dendrite Growth with Highly Concentrated Ions: Development of Surrogate Model with Ensemble Neural Networks. *Mol. Syst. Des. Eng.*, 7, 260-272.

3. **Gao, T.**, Itliong, J., Kumar, S. P., Hjorth, Z., and Nakamura, I. (2021). Polarization of ionic liquid and polymer and its implications for polymerized ionic liquids: An overview towards a new theory and simulation. *Journal of Polymer*

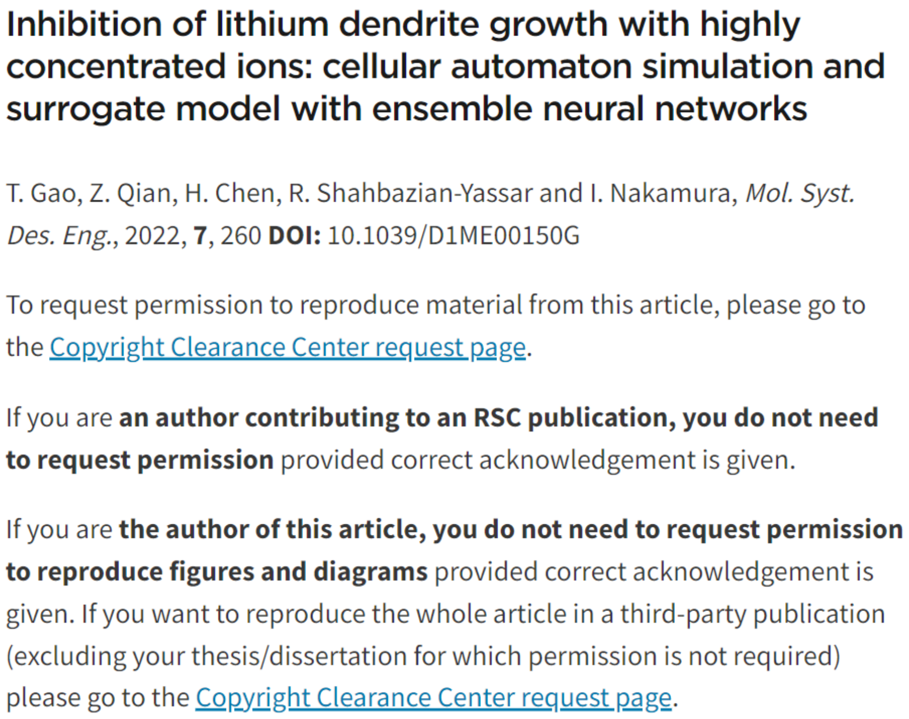
Science, 59(21), 2434.

4. Nakamura, I., Shock, C. J., Eggart, L., **Gao, T.** (2019). Theoretical Aspects of Ionic Liquids for Soft Matter Sciences. *Israel Journal of Chemistry*, 59(9), 813-823.

Appendix B

Letters of Permission

Permissions for reuse the published paper in Chapter 3, 4, 5, and 6.



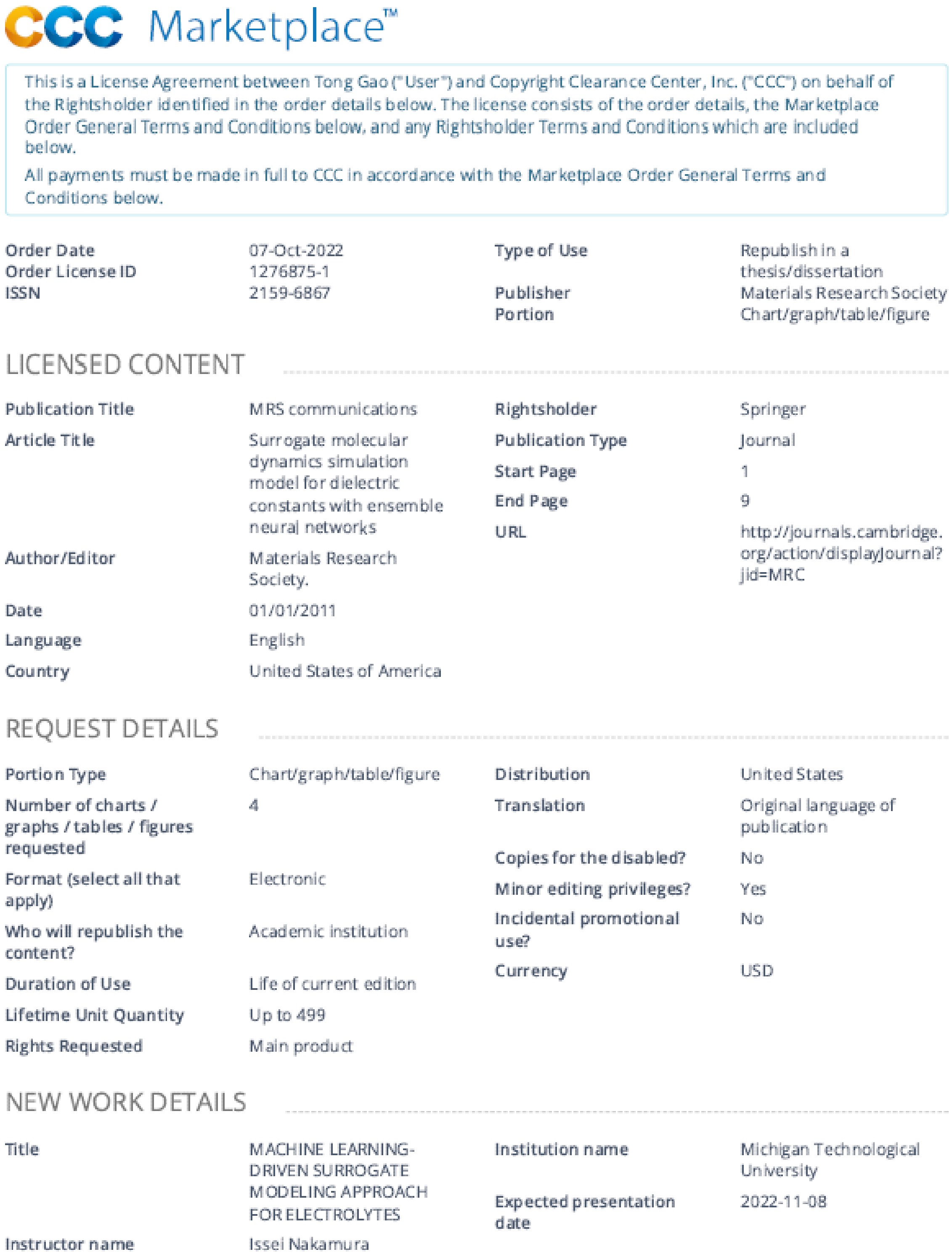
Inhibition of lithium dendrite growth with highly concentrated ions: cellular automaton simulation and surrogate model with ensemble neural networks

T. Gao, Z. Qian, H. Chen, R. Shahbazian-Yassar and I. Nakamura, *Mol. Syst. Des. Eng.*, 2022, **7**, 260 **DOI:** 10.1039/D1ME00150G

To request permission to reproduce material from this article, please go to the Copyright Clearance Center request page.

If you are **an author contributing to an RSC publication, you do not need to request permission** provided correct acknowledgement is given.

If you are **the author of this article, you do not need to request permission to reproduce figures and diagrams** provided correct acknowledgement is given. If you want to reproduce the whole article in a third-party publication (excluding your thesis/dissertation for which permission is not required) please go to the Copyright Clearance Center request page.



CCC Marketplace™

This is a License Agreement between Tong Gao ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Order General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the Marketplace Order General Terms and Conditions below.

Order Date	07-Oct-2022	**Type of Use**	Republish in a thesis/dissertation
Order License ID	1276875-1	**Publisher**	Materials Research Society
ISSN	2159-6867	**Portion**	Chart/graph/table/figure

LICENSED CONTENT

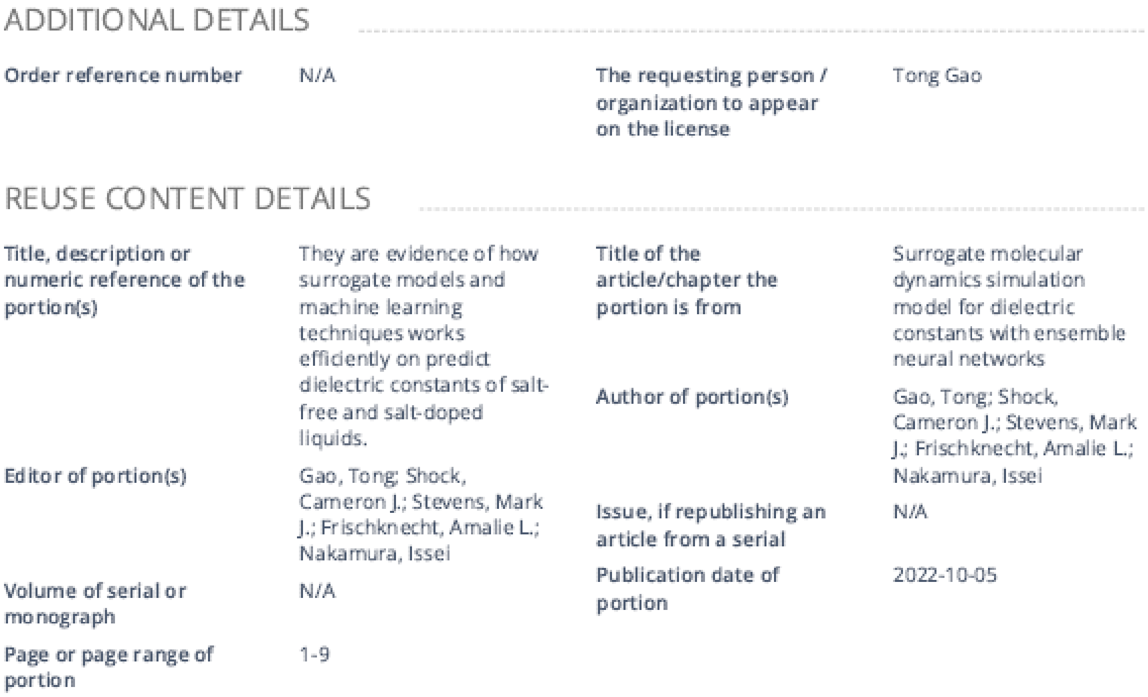
Publication Title	MRS communications	**Rightsholder**	Springer
Article Title	Surrogate molecular dynamics simulation model for dielectric constants with ensemble neural networks	**Publication Type**	Journal
		Start Page	1
		End Page	9
		URL	http://journals.cambridge.org/action/displayJournal?jid=MRC
Author/Editor	Materials Research Society.		
Date	01/01/2011		
Language	English		
Country	United States of America		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	**Distribution**	United States
Number of charts / graphs / tables / figures requested	4	**Translation**	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	**Minor editing privileges?**	Yes
Who will republish the content?	Academic institution	**Incidental promotional use?**	No
Duration of Use	Life of current edition	**Currency**	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	MACHINE LEARNING-DRIVEN SURROGATE MODELING APPROACH FOR ELECTROLYTES	**Institution name**	Michigan Technological University
		Expected presentation date	2022-11-08
Instructor name	Issei Nakamura		



ADDITIONAL DETAILS

Order reference number	N/A
The requesting person / organization to appear on the license	Tong Gao

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	They are evidence of how surrogate models and machine learning techniques works efficiently on predict dielectric constants of salt-free and salt-doped liquids.
Editor of portion(s)	Gao, Tong; Shock, Cameron J.; Stevens, Mark J.; Frischknecht, Amalie L.; Nakamura, Issei
Volume of serial or monograph	N/A
Page or page range of portion	1-9
Title of the article/chapter the portion is from	Surrogate molecular dynamics simulation model for dielectric constants with ensemble neural networks
Author of portion(s)	Gao, Tong; Shock, Cameron J.; Stevens, Mark J.; Frischknecht, Amalie L.; Nakamura, Issei
Issue, if republishing an article from a serial	N/A
Publication date of portion	2022-10-05