THE JOURNAL OF PHYSICAL CHEMISTRY B

pubs.acs.org/JPCB Article

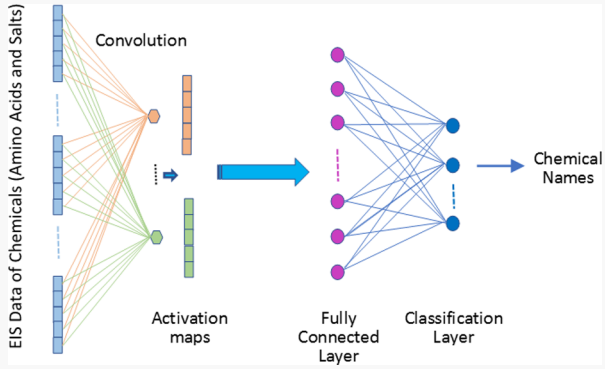
Machine Learning Analysis of the Thermodynamic Responses of In Situ Dielectric Spectroscopy Data in Amino Acids and Inorganic Electrolytes

Yong Wei,* Keith Chin,* Laura M. Barge, Scott Perl, Ninos Hermis, and Tao Wei*

Cite This: J. Phys. Chem. B 2020, 124, 11491−11500 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Dielectric spectroscopy (DS) can be a robust in situ technique for geochemical applications. In this study, we applied deep-learning techniques to DS measurement data to enable rapid science interrogation and identification of electrolyte solutions containing salts and amino acids over a wide temperature range (20 to −60 °C). For the purpose of searching for signs of life, detecting amino acids is a fundamental high priority for field and planetary instruments as amino acids are one of the building blocks for life as we know it. A convolutional neural network (CNN) with channel-wise one-dimensional filters is proposed to fulfill the task, using the DS data of amino acid and inorganic salt solutions. Experimental results show that the CNN with two convolutional layers and one fully connected layer can effectively differentiate solutions containing amino acids from those containing salts in both the liquid and solid (water ice) states. To complement the experimental measurements and CNN analysis, the diffusive behaviors of ions (K^+, Cl^-, and OH^-) were further discussed with atomistic molecular dynamics simulations performed in this work as well as the quantum simulation published in the literature. Combining DS with machine-learning techniques and simulations will greatly facilitate more real-time decision-making of mobility systems for future exploratory endeavors in other worlds beyond Earth.



1. INTRODUCTION

Electrochemistry, combined with sensitive nondestructive organic detection methods, can yield critical information about metabolic strategies, prebiotic chemistry, and biosignatures in planetary environments. The chemical and physical processes investigated by electrochemical techniques are fundamentally governed by (1) the physics of mass transport from the bulk solution to the interfaces where polarization and charge-transfer occur, for example, at a mineral interface where microbes transfer electrons to/from that solid substrate; and (2) the presence and concentration of various electroactive species in solution, for example, reduced or oxidized inorganic species or partially reduced products of carbon or other biological building blocks. Knowing this, the development of electrochemical in situ biosensors capable of dynamic characterization (including detection/identification, quantification, and selection) of basic organic or bioprecursor materials has a substantial scientific value to future space exploratory missions. In particular, dielectric spectroscopy (DS) has emerged as a powerful measurement technique for biosensor applications in characterizing organic systems containing molecules such as enzymes, proteins, and antibodies.[1−4] DS instruments are highly portable and scalable for adaptability in almost all deployment systems from large rovers to small microbots and penetrators, with measurements suitable for in situ and noninvasive capabilities using two to four probes comprised of polarizable (chemically inert) electrode materials.

From a single-frequency sweep (millihertz to megahertz), measurements from DS can provide substantial information on the chemical and physical properties of any geochemical system emanating from the molecular level. Such measurements of magnitude and phase shifts provide great details on material systems' transport dynamics and chemical energetics from electrical properties such as conductivity, dielectric constants, and relaxation time constants. Furthermore, material phases and mixtures of heterogeneous materials such as soils/water/ice also have very unique electrical properties as observed from the magnitude of impedance, capacitance, and phase angle spectra (known as Bode plots) as compared to homogeneous material phases. Traditionally, these unique

Received: October 12, 2020
Revised: November 13, 2020
Published: December 7, 2020

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

responses can be readily extracted directly from measurements using equivalent circuit models consisting of passive resistor−capacitor (R−C) combinations to provide physical interpretation of the responses and to serve as descriptors for decision trees associated with chemical/geochemical systems under scientific interrogation.[5−8] However, the incorporation of modern machine-learning techniques have significant advantages in development of pattern recognition schemes which is not yet fully realized for DS data analysis and interpretation of results.[9,10]

The detection of amino acids in extraterrestrial environments is of particular interest to the astrobiology community. For the purposes of searching for life or extant life beyond Earth, detecting organic molecules such as amino acids would be a high priority for field and planetary instruments as they are one of the building blocks for life as we know it. Amino acids have different functional groups and thus unique chemical and physical properties, making them suitable for identification and characterization by DS. Most importantly, amino acids are soluble in water, providing an excellent medium for detection by their transport properties. Certain amino acids are basic (proton acceptor) or acidic (proton donor) in nature and have charge polarity enabling detection and possible ionic selection via their electrical properties. We propose to take advantage of machine-learning techniques in order to facilitate in situ identification of amino acids and inorganic salts in samples.

Identification/classification is one of the major focuses of machine learning. Classification consists of (1) taking a feature vector x of an instance to be classified and (2) mapping the feature vector x to a class label y.[11] Classification problems are solved by finding the mapping function f: $x \rightarrow y$, which is defined as follows

$$y = f(x) \tag{1}$$

In many applications, mapping functions between features and classes are so complicated that obtaining analytic forms of the mapping functions is often impossible. Thus, a numerical representation is more desirable.

Motivated by neural science research, convolutional neural networks (CNNs) are proposed to solve the aforementioned classification problems. They are very well suited to approximating complicated and highly nonlinear mapping functions. CNN-based classifiers have had success in image classification,[12,13] including medical image classification.[14,15] CNNs can also be used on one-dimensional (1D) data classification. Examples of CNN-based classification on 1D data include using on-body sensor-collected time series to identify human activities,[16] speech recognition,[17] biological sequences classification,[18] and so forth. In the case of applying machine-learning techniques on chemical data, Sadik, et al.[19] applied support vector machines to classify organophosphate nerve agent simulants using changes in resistance versus time obtained by an array of 32 sensors. Acquarelli, et al.[20] employed CNNs to identify materials using vibrational spectroscopic data.

To solve the classification problem, a CNN is trained to approximate the mapping function between the feature vectors of instances and classes of them through supervised learning. In the mode of supervised learning, both artificial neural networks (ANNs) and CNNs take pairs of training instances, defined as (v_i, c_i), in which v_i and c_i are the feature vector and class label of the ith training instance, respectively, to adjust their parameters algorithmically. The performance of an ANN depends on the selection of input features from raw data.[21] The choice of input features is empirical, that is, trial and error. An advantage of CNNs over ANNs is that CNNs can automatically extract features from raw data. This salient characteristic is achieved by the convolution layers in a CNN.[22] More details of the data set and CNNs used in this work are discussed in Section 2.2.

Standard electrical properties spectra obtained by DS is over a wide range of frequencies in the microhertz to megahertz regions. The DS data used in this work are 1D and multichannel. Definitions of data channels are listed in Table 1. Inspired by the CNNs for image classification,[22] we propose a CNN to classify DS data to detect/classify chemicals. The proposed network has an array of channel-wise filters to automatically extract features from data obtained by DS. Extracted features are summarized at multiple abstraction levels by multiple convolutional layers before being fed into a fully connected layer for classification. Experimental results show that the proposed algorithm is effective in classifying and differentiating amino acids and inorganic salts in both the liquid and ice states.

Table 1. Definitions of Each Channel of DS Data[a]

symbol	definition	unit
Y	admittance modulus measured. $Y = \sqrt{Y_r^2 + Y_i^2}$	siemens (S)
θ	phase angle measured. $\theta = \tan^{-1}(Y_i/Y_r)$	degrees
Z	impedance modulus calculated. $Z = \frac{1}{Y}$	ohms (Ω)
Y_r	real component of the Y or in-phase portion of the measured admittance. $Y_r = Y \times \cos(\theta)$	S
Y_i	imaginary component of the Y or out-of-phase portion of the measured admittance. $Y_i = Y \times \sin(\theta)$	S
$1/\tan(\theta)$	it is associated with power dissipation factor or ratio of energy stored/energy loss	[b]
C	capacitance spectra associated with Y_i or energy stored	farads (F)
E'	relative permittivity spectra or real component of the complex dielectric spectra	[b]
E''	dielectric loss spectra (associated with energy loss because of ion motion or conductivity) or imaginary component of the complex dielectric spectra	[b]
σ	conductivity spectra of the measurement (intensive property)	S/cm
E''/E'	ratio of dielectric/dielectric loss, similar to $\tan(\theta)$ but in dielectric domain	[b]

[a]Details of these parameters are presented in ref 33. [b]* indicates it is dimensionless or unitless.

This paper is organized as follows. Section 2 describes experimental data and depicts details of the proposed CNN architecture to classify DS data in order to identify chemicals. In addition, to illustrate ions' diffusive behavior inside the ice lattice at atomistic and molecular scales, atomistic molecular dynamics (MD) simulations were performed for the two systems, both without and with applying an external electric field, also described in Section 2. Classical MD simulations with fixed atomistic charges can probe molecular structures with atomistic resolutions and sample MD behavior from subnanoseconds to microseconds.[23−32] Evaluation and dis-

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

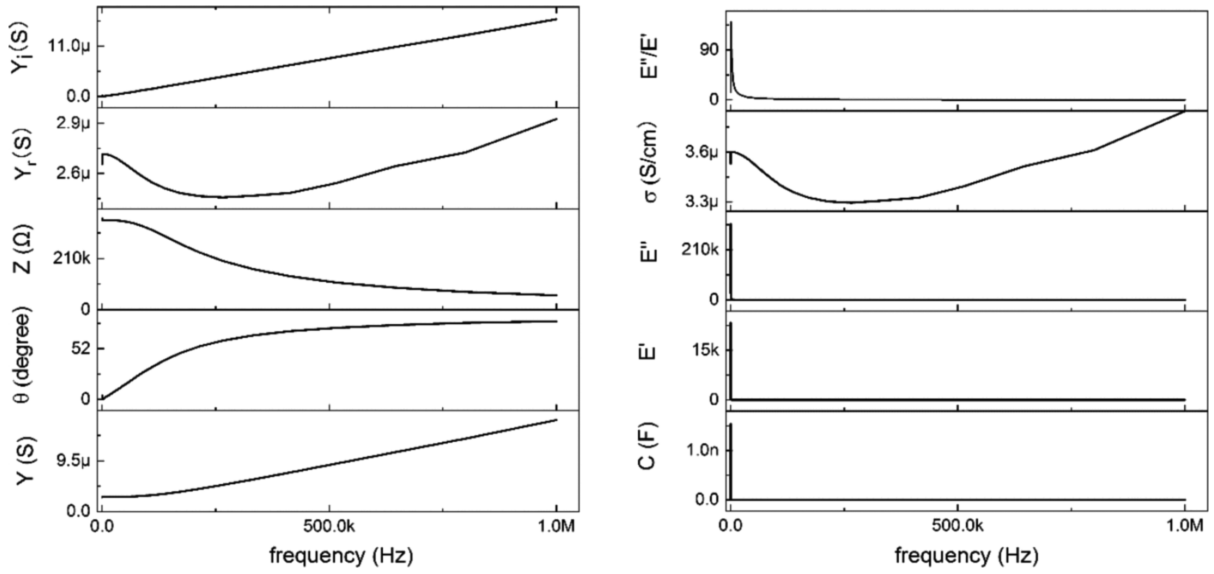


Figure 1. 1D multichannel measured and calculated DS response of glycine at 10 μm, 10 °C. A total of 10 spectra profiles over a frequency range from 20 Hz to 1 MHz were measured. The admittance modulus (Y) and phase angle (θ) were measured responses. The impedance modulus (Z), real admittance (Y_r), imaginary admittance (Y_i), real relative permittivity (E'), imaginary relative permittivity (E''), conductivity (σ), and capacitance (C) data are calculated.

cussion of experimental results are in Section 3. Finally, Section 4 concludes the paper and discusses future work.

2. METHODS

2.1. DS Measurements of Amino Acid Solutions. Amino acids were purchased as analytical grade reagents from Sigma-Aldrich (St. Louis, MO). Stock solutions of the individual amino acids were prepared at 100 mM in 18 MΩ deionized (DI) water (ELGA water purification system, PureLab-Option-Q 15BP) in 50 mL glass scintillation vials. The stock solutions were diluted further to 1 μM prior to analysis by DS. Electrolyte solutions containing KCl and KOH were prepared similarly. For this investigation, we purposely only reported results on Gly and Ala amino acid in comparison to KCl and KOH solutions in concentrations ranging from 1 μM to 100 mM for validation on the efficacy of machine-learning techniques on DS characterization; these nonpolar amino acids pose the greatest challenge to identify because of their low electrochemical activity as attributable to their chemical and physical properties.[33]

DS measurements on aqueous amino acid solutions were conducted using a novel electrochemical property cell (EPC) design. The EPC employs a 3 mm diameter stainless steel screw head to serve as the contact polarizable electrode; the physical surface area of each electrode is about 8 mm^2. The electrode configuration in the EPC is adjustable to 2 or 4 electrode measurement. The DS measurements were performed using a Wayne Kerr 6400B Precision Analyzer over a wide frequency range from 20 Hz to 1 MHz with a perturbation voltage amplitude of 100 mV with no applied voltage bias. Prior to amino acid solution measurements, the geometric factor (or cell constant), *Geo*, was accurately determined from calibration protocols using a known conductivity solution of 18 μS/cm purchased from Myron L, Inc. The *Geo* factors for EPC1 (two microelectrodes) and EPC2 (four microelectrodes) were 0.39 and 0.76 cm, respectively. Also, as part of calibration protocols, an empty cell measurement was performed to assess the system parasitic capacitance and detection limits over the same measurement conditions of voltage amplitude and frequency range. To illustrate the characteristics of the DS data, an example of the data used in our experiments is shown in Figure 1. The definitions of the eleven channels of data are given in Table 1. In Table 1, the real component Y_r and imaginary component Y_i are independent measurements of DS. Even though the other channels are calculated features, they represent various characteristics of the chemical solutions of different aspects. Thus, they are used in the CNN for the purpose of differentiating chemicals. More representative experimental measurements of DS data in amino acids and inorganic electrolytes are shown in Figures S1−S3 (in the Supporting Information document). All original experimental data can be found in the Supporting Information document.

2.2. Data Set and Architecture of the CNN with Channel-Wise Arrays of Filters. The measurement conditions of the DS data set used in training and testing are listed in Table 2. A data record consists of DS measurements (Y_r and Y_i) and the calculated features over the frequency range of 20 Hz−1 MHz. Their definitions are given in Table 1.

Table 2. DS Data Set Description[a]

chemical names	concentrations	temperatures (° C)	number of records
alanine	100 mM, 10 mM, 1 mM, 100 μM, 10 μM, 1 μM	−60, −50, −40, −30, −20, −10, 0, 10, 20	54
glycine	100 mM, 100 μM, 10 μM, 1 μM	−60, −50, −40, −30, −20, −10, 0, 10, 20	36
KCL	100 mM, 10 mM, 1 mM, 100 μM, 10 μM, 1 μM	−60, −50, −40, −30, −20, −10, 0, 10, 20	54
KOH	100 mM, 10 mM, 1 mM, 100 μM, 10 μM, 5 μM[b]	−60, −50, −40, −30, −20, −10, 0, 10, 20	53

[a]Concentrations and temperatures used in measurements. Every chemical of each concentration is measured at all the listed temperatures. [b]KOH of 5 μM is not measured at −60 °C.

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

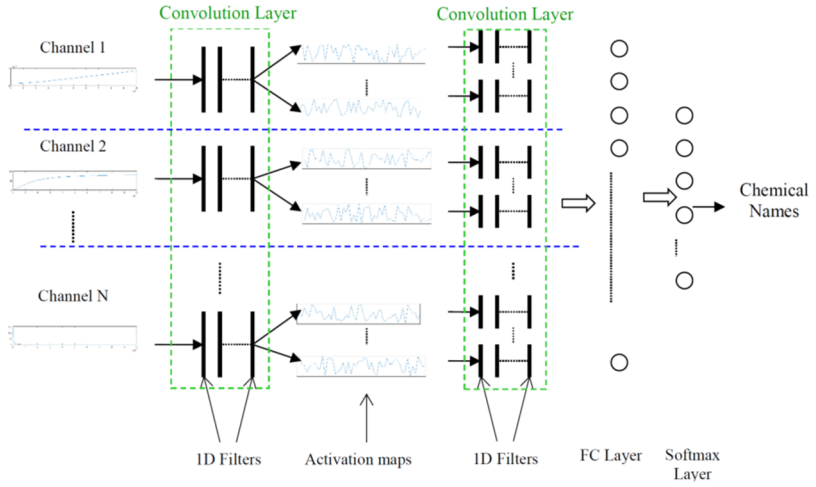


Figure 2. Architecture of the CNN-based classifier with multiple layers of channel-wise 1D filter arrays.

Each channel of the DS data represents the characteristic of chemical(s) being classified from a different aspect. Hence, channels of the DS data should be treated separately in classification. This is done via channel-wise filters in the CNN. The architecture of the proposed CNN-based DS data classifier is depicted in Figure 2. At the very front of the architecture, there are layers of channel-wise filter arrays to process the DS data at consecutive levels before sending data summaries to a fully connected layer. The outputs of the fully connected layer feed a softmax (or classification) layer to output chemical names.

The input data to the CNN in a channel are 1D, thus are filtered by a group of 1D filters in the convolutional layers. The activation map of filter f in channel c is given as the following 1D convolution (i.e., filtering) between filter $f(n,c)$ and data $g(n,c)$

$$\text{actmap}(x, c) = \sum_{n_1=-\infty}^{+\infty} g(n_1, c)\cdot f(x - n_1, c) \tag{2}$$

in which c is the channel of data. Please note that actmap(x,c) is the activation map of the cth channel of data representing characteristics of DS data of the channel. In the convolution layers, each channel is assigned a group of such 1D channel-wise filters. Filter parameters are to be learned (i.e., adjusted algorithmically) during the supervised training process. This is the source of the salient power of the CNNs by which they can automatically extract information from data being used in training. This is extremely useful when diversified categories/channels of data are used. Otherwise, without such self-learning power, input features have to be hand-tuned based on trial and error. There are multiple filters in a convolution layer, each of them will have a different set of parameters after training. Thus, a group of filters can extract various characteristics (i.e., features) from the input data for the purpose of classification.

One point worth emphasizing is that the number of parameters to be learned in a 1D filter equals to the size of the filter w_f. The total number of parameters to be learned in a convolutional layer is $w_f \times N_f \times N$, where N_f is the number of filters for a channel and N is the number of data channels. The number of parameters in a convolutional layer is usually fewer than that in a fully connected layer. This is an important feature of the proposed CNN-based classifier when only a limited amount of DS data are available in this work.

The EIS data of amino acids and inorganic salts from channels 1−11 listed in Table 1 are used to train the CNN-based classifiers. CNN-based classifiers take pairs of training instances, defined as (v_i, c_i), in which v_i and c_i are the EIS data and name of chemical of the ith training instance, respectively, to adjust their parameters (i.e., the values of filter parameters and weights of connections between the last convolutional layer and the fully connected layer, and the fully connected layer and the softmax layer, algorithmically. $v_i = [v_{i,1}, v_{i,2}, ..., v_{i,11})$, in which $v_{i,k}$ is the kth channel of the EIS data. The EIS data were collected from the experiment run, as described in Section 2.1. After training is done, given the EIS data of instances, the CNN-based classifier outputs the names of chemicals. Details about CNN-based classifier training and testing are given in Section 3.

2.3. MD Simulations. MD simulations were carried out using GROMACS package (version 4.6.5) in an isobaric−isothermal (NPT) ensemble with a periodic boundary condition. OPLS-AA forcefield parameters[34] and the four-point water model (TIP4P) were used to simulate the water/ion interactions. All simulations apply SHAKE algorithm for rigid water molecules to constrain hydrogen covalent bonds, and the integration of the dynamic equations was done using the leapfrog algorithm with a time step of 1 fs. The Berendsen barostat method was used to maintain the pressure at 1 bar through an isotropic position scaling, and Nose−Hoover thermostats were used to keep the temperature at 220 K. The particle mesh Ewald summation was utilized to calculate the long-range electrostatic interactions, with a cut-off distance of 1.6 nm. The van der Waals interactions were calculated with the cutoff at 12 Å, while the long-range dispersion effect on energy and pressure was also included.

To illustrate the diffusive behavior of ions, an ice lattice (53.9 Å × 62.2 Å × 58.7 Å) was doped with a pair of K^+ and Cl^- ions without applying an electric field. Ions were directly added into the cavity of the lattice. The simulations were carried out for 400 ns. For the system with an electric field applied, the ice lattice of the same dimensions was added with 30 ion pairs (60 ions total) by replacing water with ions, corresponding to a concentration of approximately 0.253 M, in order to simulate the conditions of briny concentrations similar

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

to those in the experimental systems. The complete time-dependent electric field pulse equation is of the form

$$E(t) = E_0 \times \cos(\omega t) \quad (3)$$

where electric field strength amplitude (E_0) is 1 V/cm and angular frequency (ω) is 1 MHz, which corresponds to a nearly constant electric field over a 1 μs simulations duration. These electric field parameters for simulation are virtually identical with the experimental conditions of the applied electric field and, therefore, offer direct comparison of simulation results to experimental observations.

3. RESULTS AND DISCUSSION

3.1. CNN Structure Used in Experiments. The CNN used for experiments consists of two convolutional layers of channel-wise 1D filters. There is an average-pooling layer following the first convolutional layer to smooth out its activation maps. A batch normalization layer is introduced following each of the convolutional layers to reduce overfitting. The fully connected layer takes the "flattened" activation maps of the second convolutional layer as its inputs. The softmax layer facilitates generating the most likely chemical names of the input instances. The architecture of the CNN used for experiments is depicted in Figure 3.

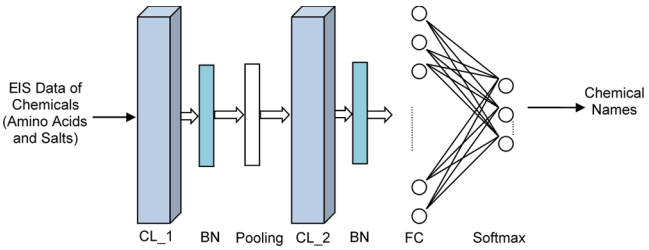


Figure 3. Architecture of the CNN used in experiments. Cl_1 and Cl_2 are convolutional layers, BN is the batch normalization layer, FC is the fully connected layer, and Softmax is the classification layer.

3.2. Performance Metrics. We use accuracy, precision, and recall to evaluate the performance of the classifiers. The performance measures are defined as follows

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

In these formulas, TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative.

3.3. Machine Learning Experiments and Evaluation. The system is implemented in Matlab 2019b. All experiments are conducted in CentOS Linux 7 (Core) on a node of 8 NVIDIA Volta V100 GPUs, 16 GB/GPU memory, and 192 GB RAM. In our experiments, the CNN hyperparameters and the range of their values used in parameter sweep are as follows:

- the size of the 1D filters in the convolutional layers $w \times 1$, in which $w \in \{4, 8, 12, 16, 20\}$.
- the number of filters in the convolutional layers $N_f \in \{10, 20, 30, 40, 50\}$. The first and second convolutional layers have the same filter size and number of filters.
- the size of the average pooling window = 2
- the stride of all layers = 1.
- the number of neurons in the fully connected layer $\in \{50, 100, 150, 200, 250, 300, 350, 400, 450, 500\}$
- the number of outputs of the softmax classification layer = the number of chemical classes in experiments

Zero padding is done in a way that the output of a layer has the same size as the input. CNN parameters are updated using the stochastic gradient descent with momentum algorithm with 0.5 momentum. The learning rate is 0.001. The maximum number of epochs used in training is 800. The minibatch size is 128. Training and validation data are shuffled once before training. To determine optimal combinations of the aforementioned CNN hyperparameters, a CNN parameter combination sweep is performed to find the best CNNs which yield the highest classification accuracy, defined by eq 4. Then, the CNNs with the best hyperparameter combinations are trained and tested multiple times using shuffled data for both training and testing. In both the parameter sweep and the simulation steps, 80% randomly selected data are used to train the CNNs and the rest 20% are used to test their performance. Because the ranges of features in the electrical spectra vary significantly, data values are normalized to the range of $[-1, +1]$ before they are fed into CNNs for both training and testing purposes.

In our experiments, the CNNs with the best hyperparameter combinations chosen by parameter sweep are trained and tested 1000 times in total using shuffled data for both training and testing. We use the metrics defined in eqs 4–6 to measure the performance of the CNN-based classifiers. The mean and standard deviations of these metrics are listed in the tables in this section. Tables 3, 5, and 7 display classification accuracies of various classifiers using the positive temperatures, negative excluding −10 °C, and negative including −10 °C data, respectively. These solutions at −10 °C are in nonequilibrium conditions during measurement, thus have characteristics of both liquid and ice phases. The precision and recall of classification are listed in Tables 4, 6, and 8. The classification accuracies are plotted in Figure 4 for the convenience of comparison.

Table 3. Accuracy of Classifiers[a]

classifiers	mean accuracy (%)	accuracy std (%)
alanine−KOH	87	13
glycine−KOH	78	19
alanine−KCl	85	13
glycine−KCl	81	16
alanine−glycine	72	19
KOH−KCl	80	19
alanine−glycine−KOH	63	13
{alanine, glycine}[b]−KOH	86	11
alanine−glycine−KCl	61	14
{alanine, glycine}[b]−KCl	85	12
alanine−glycine−KOH−KCl	62	13
{alanine, glycine}[b]−{KOH, KCl}[b]	82	11

[a]Positive temperatures. [b]{X, Y} means chemicals X and Y are treated as one class.

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

Table 4. Precision (P) and Recall (R) of Classifiers[a]

classifiers	alanine	glycine	KOH	KCl	{alanine, glycine}[b]	{KOH, KCl}[b]
alanine−KOH	P = 85, R = 94	NA	P = 95, R = 80	NA	NA	NA
glycine−KOH	NA	P = 71, R = 78	P = 87, R = 77	NA	NA	NA
alanine−KCl	P = 86, R = 89	NA	NA	P = 91, R = 82	NA	NA
glycine−KCl	NA	P = 78, R = 78	NA	P = 87, R = 84	NA	NA
alanine−glycine	P = 79, R = 76	P = 64, R = 64	NA	NA	NA	NA
KOH−KCl	NA	NA	P = 80, R = 82	P = 82, R = 77	NA	NA
alanine−glycine−KOH	P = 59, R = 81	P = 22, R = 17	P = 83, R = 75	NA	NA	NA
{alanine, glycine}[b]−KOH	NA	NA	P = 87, R = 75	NA	P = 83, R = 93	NA
alanine−glycine−KCl	P = 56, R = 68	P = 28, R = 25	NA	P = 82, R = 77	NA	NA
{alanine, glycine}[b]−KCl	NA	NA	NA	P = 83, R = 79	P = 89, R = 88	NA
alanine−glycine−KOH−KCl	P = 54, R = 81	P = 19, R = 16	P = 78, R = 68	P = 83, R = 68	NA	NA
{alanine, glycine}[b]−{KOH, KCl}[b]	NA	NA	NA	NA	P = 78, R = 88	P = 88, R = 78

[a]Positive temperatures. [b]{X, Y} means chemicals X and Y are treated as one class. P = precision (%). R = recall (%).

Table 5. Accuracy of Classification[a]

classifiers	mean accuracy (%)	accuracy std (%)
alanine−KOH	67	14
glycine−KOH	95	9
alanine−KCl	84	12
glycine−KCl	93	11
alanine−glycine	62	17
KOH−KCl	91	10
alanine−glycine−KOH	55	13
{alanine, glycine}[b]−KOH	73	11
alanine−glycine−KCl	66	12
{alanine, glycine}[b]−KCl	82	10
alanine−glycine−KOH−KCl	60	10
{alanine, glycine}[b]−{KOH, KCl}[b]	84	7

[a]Negative temperatures, excluding −10 °C data. [b]{X, Y} means chemicals X and Y are treated as one class.

Table 7. Accuracy of Classification[a]

classifiers	mean accuracy (%)	accuracy std (%)
alanine−KOH	51	13
glycine−KOH	61	15
alanine−KCl	77	12
glycine−KCl	74	14
alanine−glycine	57	15
KOH−KCl	76	11
alanine−glycine−KOH	38	11
{alanine, glycine}[b]−KOH	53	12
alanine−glycine−KCl	55	11
{alanine, glycine}[b]−KCl	78	10
alanine−glycine−KOH−KCl	43	9
{alanine, glycine}[b]−{KOH, KCl}[b]	64	10

[a]Negative temperatures, including −10 °C data. [b]{X, Y} means chemicals X and Y are treated as one class.

Tables 3 and 4 list the classification performance of chemicals in the liquid state. Tables 5 and 6 show the classification performance of chemicals in the solid state. Here, we explain the reasons why classification is done in the liquid and solid states separately. Fundamentally, dielectric (or impedance) spectroscopy can enable measurements that separate molecular processes on the basis of response time, providing a unique relaxation frequency along with a signature variation with frequency (10^{-6} to 10^{12} Hz).[35] The relaxation phenomenon is a time-dependent process (i.e., reaction, diffusion, etc.) in response to an applied electric field resulting in a dielectric displacement of charge carriers within a material. For bounded charges, relaxation is related to fluctuations of dipoles in molecules or part of them. For free charge carriers, displacement leads to translational (drift) motion causing conduction contribution to dielectric response.

The dielectric displacement, $D(t)$, is a consequence of a phase shift, which is the time difference between the material response to outer applied E-field, $[E(t)]$. In real chemical systems, the phase shift can have values ranging from 0 to 90°, where the 0° phase angle indicates that responses are dominated by free charge carriers inducing translational

Table 6. Precision (P) and Recall (R) of Classifications[a]

classifiers	alanine	glycine	KOH	KCl	{alanine, glycine}[b]	{KOH, KCl}[b]
alanine−KOH	P = 69, R = 54	NA	P = 69, R = 77	NA	NA	NA
glycine−KOH	NA	P = 93, R = 91	P = 97, R = 97	NA	NA	NA
alanine−KCl	P = 83, R = 83	NA	NA	P = 89, R = 86	NA	NA
glycine−KCl	NA	P = 88, R = 88	NA	P = 96, R = 95	NA	NA
alanine−glycine	P = 73, R = 71	P = 39, R = 44	NA	NA	NA	NA
KOH−KCl	NA	NA	P = 92, R = 89	P = 92, R = 92	NA	NA
alanine−glycine−KOH	P = 36, R = 30	P = 44, R = 48	P = 70, R = 79	NA	NA	NA
{alanine, glycine}[b]−KOH	NA	NA	P = 90, R = 75	NA	P = 80, R = 90	NA
alanine−glycine−KCl	P = 55, R = 57	P = 33, R = 34	NA	P = 84, R = 82	NA	NA
{alanine, glycine}[b]−KCl	NA	NA	NA	P = 90, R = 75	P = 80, R = 90	NA
alanine−glycine−KOH−KCl	P = 33, R = 27	P = 38, R = 41	P = 61, R = 68	P = 82, R = 81	NA	NA
{alanine, glycine}[b]−{KOH, KCl}[b]	NA	NA	NA	NA	P = 88, R = 66	P = 84, R = 94

[a]Negative temperatures, excluding −10 °C data. [b]{X, Y} means chemicals X and Y are treated as one class. P = precision (%). R = recall (%).

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

Table 8. Precision (P) and Recall (R) of Classification[a]

classifiers	alanine	glycine	KOH	KCl	{alanine, glycine}[b]	{KOH, KCl}[b]
alanine−KOH	P = 43, R = 48	NA	P = 59, R = 54	NA	NA	NA
glycine−KOH	NA	P = 34, R = 35	P = 72, R = 72	NA	NA	NA
alanine−KCl	P = 73, R = 77	NA	NA	P = 84, R = 77	NA	NA
glycine−KCl	NA	P = 58, R = 61	NA	P = 84, R = 79	NA	NA
alanine−glycine	P = 65, R = 69	P = 39, R = 36	NA	NA	NA	NA
KOH−KCl	NA	NA	P = 77, R = 77	P = 78, R = 75	NA	NA
alanine−glycine−KOH	P = 37, R = 41	P = 17, R = 14	P = 47, R = 47	NA	NA	NA
{alanine, glycine}[b]−KOH	NA	NA	P = 51, R = 34	NA	P = 55, R = 70	NA
alanine−glycine−KCl	P = 50, R = 53	P = 15, R = 15	NA	P = 77, R = 73	NA	NA
{alanine, glycine}[b]−KCl	NA	NA	NA	P = 84, R = 69	P = 77, R = 87	NA
alanine−glycine−KOH−KCl	P = 36, R = 35	P = 8, R = 6	P = 39, R = 41	P = 64, R = R = 67	NA	NA
{alanine, glycine}[b]−{KOH, KCl}[b]	NA	NA	NA	NA	P = 51, R = 49	P = 72, R = 72

[a]Negative temperatures, including −10 °C data. [b]{X, Y} means chemicals X and Y are treated as one class. P = precision (%). R = recall (%).

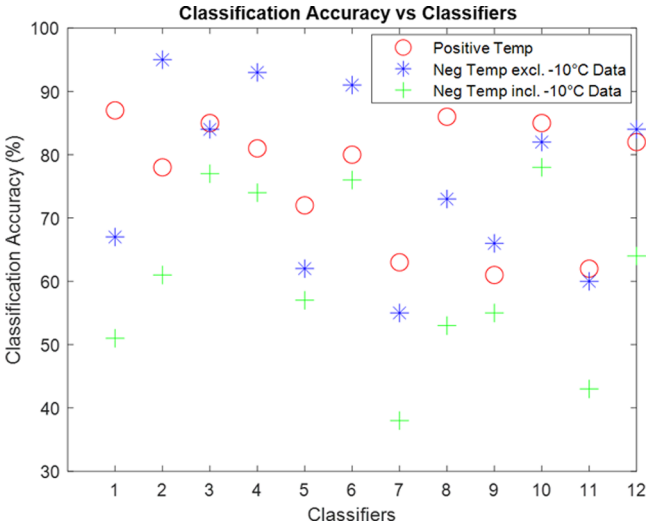


Figure 4. Classification accuracy versus classifiers. Classifiers: 1 = alanine−KOH, 2 = glycine−KOH, 3 = alanine−KCl, 4 = glycine−KCl, 5 = alanine−glycine, 6 = KOH−KCl, 7 = alanine−glycine−KOH, 8 = {alanine, glycine}*−KOH, 9 = alanine−glycine−KCl, 10 = {alanine, glycine}*−KCl, 11 = alanine−glycine−KOH−KCl, and 12 = {alanine, glycine}*−{KOH, KCl}*. *Note: {X, Y} means chemicals X and Y are treated as one class.

motion and can be modeled by resistors. At 90° phase angle, one can say it is because of bonded charge carriers similar to behavior in capacitors. In real chemical systems responding to electrical spectroscopy, it is usually a combination of RC units which provides the relaxation times of many physical/chemical processes that can be detected over several frequency decades using this technique. Therefore, each of the parameters listed below has some information of the behavior of free versus bonded charge carriers; hence, complex numbers are used to separate contributions of real (in-phase) versus imaginary (out-of-phase) components of the spectrum. The phase angle is just the ratio of real to imaginary values.

In the liquid state, free charge carriers dominate the response; it is directly related to important chemical properties including ionic solubility, concentration, transport properties (i.e., mobility and diffusion), and interfacial double-layer charging. In the ice state, the bonded charge from ice dominates the response; these properties include relaxation time constants, static (low frequency) relative permittivity, and defect or ion transport. Because of these reasons, it is reasonable to identify chemicals in liquid and ice states separately.

Tables 7 and 8 display the classification performance of chemicals at negative temperatures, including data of −10 °C. As beforementioned, these solutions are two-phase systems at −10 °C, thus have characteristics of both liquid and ice exhibiting nonequilibrium conditions during measurement. Hence, excluding data of −10 °C used in classification results in better accuracies for negative temperatures, as shown in Figure 4.

The accuracies are higher when differentiating amino acids {alanine and glycine} from inorganic salts {KCl and KOH} than those differentiating among amino acids. As shown in Figure 4, classifiers to differentiate amino acids from inorganic salts (classifiers 1, 2, 3, 4, 8, 10, and 12) yield better accuracies than classifiers to differentiate amino acids from each other (classifiers 5, 7, 9, and 11), in both liquid and ice states. There are several observed responses that best contribute to this result. In the liquid state, inorganic salts have substantially higher ionic strength than those of amino acids. Consequently, both the responses in relative permittivity and conductivity spectra exhibited unique features; the relaxation times in the relative permittivity spectra are lower, and the magnitudes in the conductivity spectra are higher for inorganic salts. Similarly, in ice states, their electrical properties in ice are strongly impacted by charge density, concentration, and physical properties of the solvated species (i.e., molecular size).[36] Therefore, their relaxation times are lower and conductivity is higher for inorganic salts despite substantial lower magnitudes over the entire frequency range as compared to the liquid state, which is consistent with previous studies performed by Kawada and others.[37−39] Because of these differences between amino acids and inorganic salt DS measurements, it is easier for the classifiers to differentiate amino acids from salts than to differentiate amino acids from each other.

The case of differentiating KCl from KOH is worth more attention. The classification accuracies are 80% (in Table 3) and 91% (in Table 5) in the liquid and ice states, respectively. These results in electrical properties are mainly attributed to the difference in their transport or diffusive behavior. In ices, the doped ions cause Bjerrum defects in the lattice that enhances the system's static conductivity.[40]

3.4. MD Simulations. To complement the experimental measurement and machine learning analysis, we performed atomistic MD simulations in the ice I_h. Figure 5 shows our

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

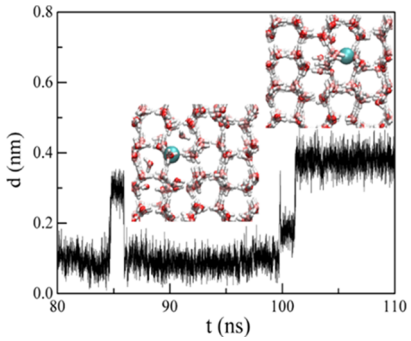


Figure 5. Displacement (d) of a K^+ ion inside ice simulated with atomistic MD simulations at temperature of 220 K and 1 bar. The initial position is used as the reference to compute the displacement. Two snapshots display the configurations before and after the hopping showing the displacement trajectory of the ion throughout the simulation.

simulation results of a single pair of K^+ and Cl^- ions doped directly inside the lattice cavity inside the lattice of ice I_h at 220 K without applying an electric field. Our simulations show that K^+ and Cl^- ions can be preferentially doped into the ice lattice, causing local distortion of the ice lattice. Moreover, they exhibit hopping behavior (Figure 5) driven by thermal motions similar to other semiconducting materials.[41] As shown in Table 9, three simulations were also carried out in an external electric field using three different initial configurations, where the K^+ and Cl^- ions were doped inside the ice lattice by replacing water molecules with ions at different crystal sites. By monitoring the temporal profile of ions, the jump count of both K^+ and Cl^- ions from those three simulations exhibited excellent repeatability (Table 9), which indicates that jump frequencies are relatively independent on initial ion locations. This is attributable to the slow diffusion processes in solid-state materials. The hop frequency of K^+ ions is about five times that of Cl^- ions at 200−220 K, and the increase in temperature enhances ions' hopping rates (Table 9). The variance in the hopping rates of such simple electrolytes (K^+ and Cl^-) is highly dependent on factors including ions' charges, ionic radii, surface charge densities, water-ion van der Waals interactions, and the masses. The ionic radius of Cl^- is 167 pm, which is larger than that of K^+ at 152 pm, which can lead to the difference in surface charge densities and diffusivities.

Table 9. K^+ and Cl^- Ions' Hopping Rate of Three Simulation Cases Using Different Initial Configurations Over the Simulation Time Range of 200 ns to 1 μs

jump rate (jumps/μs) and jump rate ratios	simulation 1	simulation 2	simulation 3	average
K^+ (200 K)	61.9	67.4	81.1	70.1
Cl^- (200 K)	14.0	18.7	12.1	14.9
K^+ (220 K)		214.5		
Cl^- (220 K)		90.5		
ratio jump rate K^+/Cl^-	4.4	3.6	6.7	4.9
ratio jump rate K^+ 220 K/200 K		3.2		
ratio jump rate Cl^- 220 K/200 K		4.8		

As shown in the experimental measurement (Figures S1−S3 in the Supporting Information), in comparison with KOH-doped ice systems, KCl-doped ices have a higher conductivity at the low concentration (1 μM and 1 mM) at 200−250 K. However, at the high concentration (100 mM), the conductivity of the KCl-doped ice is lower than that doped with KOH. The likely explanation is found in the more complicated transport mechanism of hydroxide (OH^-) ions as compared to K^+ and Cl^- ions in association with a combinations of factors such as coordination state, proton transfer, nuclear quantum effects, and solvation shell effects.[42−44] Quantum dynamics simulation indicated the possible presence of "structural diffusion", where charge migration occurs through the water hydration bond network.[42,45] However, the transport of OH^- ions can be more complicated because they can adopt different dynamic structural configurations.[42,46] An OH^- can be present with in-the-lattice bulk configuration, where OH^- accepts three hydrogen bonds and donates one to its neighboring water molecules.[42,47] OH^- can also adopt off-the-lattice configuration, where an OH^- anion accepts four hydrogen bonds at the hydroxide oxygen site and the OH^- axis points into a cavity in the crystal structure, when the temperature decreases, particularly below the ice melting temperature.[42,47,48] Such a unique hyper-coordinated state slows down the rotational motions and decreases the structural diffusivity of OH^- in the ice lattice.[42,48] In water, the stable nonplanar and hyper-coordinated solvation structure of OH^- also suppresses proton transfer and slows down its diffusion compared to H_3O^+.[49] At a microscopic level, ion concentration can also play a critical role in ice transport properties as it can lead to localized phase change behavior. Both our experimental measurement of the conductivity and other studies using techniques such as nuclear magnetic resonance[50] suggested that the liquid-like phase often coexists with the ice crystal when the concentration of KOH is more than 10 mM. Further studies of hydroxide diffusive behavior in the ice at different concentrations will be conducted in our future works to provide a more comprehensive explanation of the experimental data.

4. CONCLUSIONS

DS measurement responses can provide critical electrical properties based on the unique response times of processes such as molecular transport and relaxation phenomena over a wide frequency range (10^{-6} to 10^6 Hz) attributed to the presence of free or bound charge carriers. As an alternative to equivalent circuit modeling in extracting important physical interpretation of the measured response, we applied machine-learning algorithms in the form of convolution neural networks in an attempt to identify amino acid organic molecules and inorganic salts solution despite similarities in their spectroscopic responses over a temperature range of 20 to −60 °C. As this work has demonstrated, DS data when characterized by such machine-learning techniques can be used to differentiate between amino acids and salts, and to some extent between different amino acids, for the case of alanine and glycine in salt solutions containing K^+, OH^-, and Cl^- ions.

A CNN with two convolutional and one fully connected layer is proposed in this work. The convolutional layers in the CNN consist of channel-wise 1D filters to extract information from the DS data of amino acid and inorganic salt solutions. Information of data is extracted and summarized at two consecutive levels by channel-wise 1D filters before the summary is fed into the fully connected layer for identification. Experimental results show that classifiers based on the

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500

proposed CNN structure can effectively differentiate amino acid solutions from salt solutions in both the liquid and solid states. The proposed DS identifying approach will greatly facilitate more real-time decision-making of mobility systems for future exploratory endeavors in other worlds beyond Earth.

Classical atomistic MD simulations demonstrated microscopic structure and diffusive behavior of ionic ions inside the ice lattice. Simulations showed that K^+ cations display more frequent hopping frequencies than Cl^- anions, whether an electric field is applied or not. As shown in the quantum simulation in the previous studies in the literature,[47,48] compared to point charge ions (K^+ and Cl^-), the diffusion of OH^- ions is more complex, mainly because of the various hyper-coordinated states.

In the future, we will apply the CNN-based identification method to more amino acids and salts. Moreover, to address the issue of identifying present or past habitability of soil and ice environments, we plan to extend the application of the CNN-based identification to solutions containing a mixture of amino acids and salts and solutions containing just salts. Further studies at quantum levels will also be conducted to investigate the correlation between the hydroxide ions' diffusivity and their concentration in the ice lattice. Further simulations at the quantum and atomistic scale will also be performed to study relaxation behavior of amino acids inside ice and water in an alternating electric field to compare with the behavior of electrolytes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c09266.

All original experimental measurement data (ZIP)

Representative experimental measurements of DS data in amino acids and inorganic electrolytes (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Yong Wei − *Department of Computer Science and Information Systems, University of North Georgia, Dahlonega, Georgia 30597, United States*; Email: yong.wei@ung.edu

Keith Chin − *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, United States*; Email: keith.b.chin@jpl.nasa.gov

Tao Wei − *Chemical Engineering Department, Howard University, Washington, D.C. 20059, United States*; orcid.org/0000-0001-6888-1658; Email: tao.wei@howard.edu

Authors

Laura M. Barge − *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, United States*; orcid.org/0000-0002-2187-540X

Scott Perl − *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, United States*

Ninos Hermis − *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.0c09266

Author Contributions

Conceptualization: Y.W., K.C., and T.W.; Methodology: Y.W., K.C., and T.W.; Software: Y.W. and T.W.; Formal Analysis: T.W.; Resources: K.C., L.M.B., S.P., and N. H.; Writing—Original Draft: Y.W., K.C., and T.W.; Writing—Review and Editing; L.M.B., S.P., and N.H. All authors have confirmed the manuscript and given approval to the final version of it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research was partially carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. Specifically, it used the Bridges system, which is supported by NSF award number ACI-1445606, at the Pittsburgh Supercomputing Center (PSC). The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC, visualization, database, or grid resources that have contributed to the research results reported within this paper. T.W. thanks the support of the National Science Foundation grant (1943999).

■ REFERENCES

(1) Pänke, O.; Balkenhohl, T.; Kafka, J.; Schäfer, D.; Lisdat, F. *Impedance Spectroscopy and Biosensing for the 21st Century*; Springer, 2007.

(2) Tertiş, M.; Florea, A.; Feier, B.; Marian, I. O.; Silaghi-Dumitrescu, L.; Cristea, A.; Săndulescu, R.; Cristea, C. Electrochemical Impedance Studies on Single and Multi-Walled Carbon Nanotubes—Polymer Nanocomposites for Biosensors Development. *J. Nanosci. Nanotechnol.* **2015**, *15*, 3385−3393.

(3) Ohno, R.; Ohnuki, H.; Wang, H.; Yokoyama, T.; Endo, H.; Tsuya, D.; Izumi, M. Electrochemical impedance spectroscopy biosensor with interdigitated electrode for detection of human immunoglobulin A. *Biosens. Bioelectron.* **2013**, *40*, 422−426.

(4) Hu, Y.; Zuo, P.; Ye, B.-C. Label-free electrochemical impedance spectroscopy biosensor for direct detection of cancer cells based on the interaction between carbohydrate and lectin. *Biosens. Bioelectron.* **2013**, *43*, 79−83.

(5) Boukamp, B. A. A linear Kronig-Kramers transform test for immittance data validation. *J. Electrochem. Soc.* **1995**, *142*, 1885.

(6) K'Owino, I. O.; Sadik, O. A. Impedance spectroscopy: a powerful tool for rapid biomolecular screening and cell culture monitoring. *Electroanalysis* **2005**, *17*, 2101−2113.

(7) Han, P.-j.; Zhang, Y.-f.; Chen, F. Y.; Bai, X.-h. Interpretation of electrochemical impedance spectroscopy (EIS) circuit model for soils. *J. Cent. South Univ.* **2015**, *22*, 4318−4328.

(8) Bonanos, N.; Pissis, P.; Macdonald, J. R. Impedance spectroscopy of dielectrics and electronic conductors. In *Characterization of Materials*; Kaufmann, E. N., Ed.; John Wiley & Sons, Inc.: 2012; pp 1−14.

(9) Macdonald, D. D. *Electrochemical and Optical Techniques for the Study and Monitoring of Metallic Corrosion*; Springer, 1991; pp 31−68.

(10) Macdonald, D. D. Reflections on the history of electrochemical impedance spectroscopy. *Electrochim. Acta* **2006**, *51*, 1376−1388.

(11) Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J. A. W. M.; Van Ginneken, B.; Sánchez, C. I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60−88.

(12) Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet

large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211−252.
(13) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012; pp 1097−1105.
(14) Tajbakhsh, N.; Shin, J. Y.; Gurudu, S. R.; Hurst, R. T.; Kendall, C. B.; Gotway, M. B.; Liang, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imag.* **2016**, *35*, 1299−1312.
(15) Payan, A.; Montana, G. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. **2015**, arXiv:1502.02506. arXiv preprint.
(16) Yang, J.; Nguyen, M. N.; San, P. P.; Li, X.; Krishnaswamy, S. Deep convolutional neural networks on multichannel time series for human activity recognition. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, 2015; pp 3995−4001.
(17) Abdel-Hamid, O.; Mohamed, A.-r.; Jiang, H.; Penn, G. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012; pp 4277−4280.
(18) Alipanahi, B.; Delong, A.; Weirauch, M. T.; Frey, B. J. Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. *Nat. Biotechnol.* **2015**, *33*, 831−838.
(19) Sadik, O.; Land, W. H.; Wanekaya, A. K.; Uematsu, M.; Embrechts, M. J.; Wong, L.; Leibensperger, D.; Volykin, A. Detection and classification of organophosphate nerve agent simulants using support vector machines with multiarray sensors. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 499−507.
(20) Acquarelli, J.; van Laarhoven, T.; Gerretzen, J.; Tran, T. N.; Buydens, L. M. C.; Marchiori, E. Convolutional neural networks for vibrational spectroscopic data analysis. *Anal. Chim. Acta* **2017**, *954*, 22−31.
(21) May, R.; Dandy, G.; Maier, H. Review of input variable selection methods for artificial neural networks. *Artificial Neural Networks-Methodological Advances and Biomedical Applications*, 2011; Vol. *10*, p 16004.
(22) Zeiler, M. D.; Fergus, R. Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, 2014; pp 818−833.
(23) Wei, T.; Huang, T.; Qiao, B.; Zhang, M.; Ma, H.; Zhang, L. Structures, dynamics, and water permeation free energy across bilayers of lipid A and its analog studied with molecular dynamics simulation. *J. Phys. Chem. B* **2014**, *118*, 13202−13209.
(24) Wei, T.; Zhang, L.; Zhao, H.; Ma, H.; Sajib, M. S. J.; Jiang, H.; Murad, S. Aromatic polyamide reverse-osmosis membrane: an atomistic molecular dynamics simulation. *J. Phys. Chem. B* **2016**, *120*, 10311−10318.
(25) Zhang, T.; Wei, T.; Han, Y.; Ma, H.; Samieegohar, M.; Chen, P.-W.; Lian, I.; Lo, Y.-H. Protein−ligand interaction detection with a novel method of transient induced molecular electronic spectroscopy (TIMES): experimental and theoretical studies. *ACS Cent. Sci.* **2016**, *2*, 834−842.
(26) Sajib, M. S. J.; Wei, Y.; Mishra, A.; Zhang, L.; Nomura, K.-i.; Kalia, R. K.; Vashishta, P.; Nakano, A.; Murad, S.; Wei, T. Atomistic Simulations of Biofouling and Molecular Transfer of Crosslinked Aromatic Polyamide Membrane for Desalination. *Langmuir* **2020**, *36*, 7658.
(27) Wei, T.; Carignano, M. A.; Szleifer, I. Lysozyme adsorption on polyethylene surfaces: why are long simulations needed? *Langmuir* **2011**, *27*, 12074−12081.
(28) Nakano, C. M.; Byun, H. S.; Ma, H.; Wei, T.; El-Naggar, M. Y. A framework for stochastic simulations and visualization of biological electron-transfer dynamics. *Comput. Phys. Commun.* **2015**, *193*, 1−9.
(29) Wei, T.; Mu, S.; Nakano, A.; Shing, K. A hybrid multi-loop genetic-algorithm/simplex/spatial-grid method for locating the optimum orientation of an adsorbed protein on a solid surface. *Comput. Phys. Commun.* **2009**, *180*, 669−674.
(30) Wei, T.; Ma, H.; Nakano, A. Decaheme cytochrome MtrF adsorption and electron transfer on gold surface. *J. Phys. Chem. Lett.* **2016**, *7*, 929−936.
(31) Wei, T.; Sajib, M. S. J.; Samieegohar, M.; Ma, H.; Shing, K. Self-Assembled Monolayers of an Azobenzene Derivative on Silica and Their Interactions with Lysozyme. *Langmuir* **2015**, *31*, 13543−13552.
(32) Wei, T.; Carignano, M. A.; Szleifer, I. Molecular dynamics simulation of lysozyme adsorption/desorption on hydrophobic surfaces. *J. Phys. Chem. B* **2012**, *116*, 10189−10194.
(33) Chin, K. B.; Chi, I.; Pasalic, J.; Huang, C.-K.; Barge, L. M. An introductory study using impedance spectroscopy technique with polarizable microelectrode for amino acids characterization. *Rev. Sci. Instrum.* **2018**, *89*, 045108.
(34) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(35) Kremer, F.; Schönhals, A. *Broadband Dielectric Spectroscopy*; Springer Science & Business Media, 2002; pp 35−56.
(36) Petrenko, V. F.; Whitworth, R. W. *Physics of Ice*; OUP Oxford, 1999.
(37) Kawada, S. Dielectric anisotropy in ice Ih. *J. Phys. Soc. Jpn.* **1978**, *44*, 1881−1886.
(38) Kawada, S. Dielectric properties of KOH-doped D2O ice. *J. Phys. Soc. Jpn.* **1989**, *58*, 295−300.
(39) Kawada, S.; Ru, R. G.; Abo, M. Dielectric Properties and 110 K Anomalies in KOH- and HCl-Doped Ice Single Crystals. *J. Phys. Chem. B* **1997**, *101*, 6223−6225.
(40) Conde, M. M.; Rovere, M.; Gallo, P. Spontaneous NaCl-doped ice at seawater conditions: focus on the mechanisms of ion inclusion. *Phys. Chem. Chem. Phys.* **2017**, *19*, 9566−9574.
(41) Warren, M.; Rottler, J. Atomistic mechanism of physical ageing in glassy materials. *EPL* **2009**, *88*, 58005.
(42) Ma, Z.; Tuckerman, M. E. On the connection between proton transport, structural diffusion, and reorientation of the hydrated hydroxide ion as a function of temperature. *Chem. Phys. Lett.* **2011**, *511*, 177−182.
(43) Codorniu-Hernández, E.; Kusalik, P. G. Hydroxyl radicals in ice: insights into local structure and dynamics. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11639−11650.
(44) Codorniu-Hernández, E.; Kusalik, P. G. Insights into the solvation and mobility of the hydroxyl radical in aqueous solution. *J. Chem. Theory Comput.* **2011**, *7*, 3725−3732.
(45) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. Ab initio molecular dynamics simulation of the solvation and transport of H3O+ and OH-ions in water. *J. Phys. Chem.* **1995**, *99*, 5749−5752.
(46) Chandra, A.; Tuckerman, M. E.; Marx, D. Connecting solvation shell structure to proton transport kinetics in hydrogen−bonded networks via population correlation functions. *Phys. Rev. Lett.* **2007**, *99*, 145901.
(47) Cwiklik, L.; Devlin, J. P.; Buch, V. Hydroxide impurity in ice. *J. Phys. Chem. A* **2009**, *113*, 7482−7490.
(48) Cwiklik, L.; Buch, V. Hydroxide trapped in the interior of ice: A computational study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1294−1296.
(49) Chen, M.; Zheng, L.; Santra, B.; Ko, H.-Y.; DiStasio, R. A., Jr.; Klein, M. L.; Car, R.; Wu, X. Hydroxide diffuses slower than hydronium in water because its solvated structure inhibits correlated proton transfer. *Nat. Chem.* **2018**, *10*, 413−419.
(50) Schildmann, S.; Schöler, A.; Nowaczyk, A.; Geil, B.; Böhmer, R. Salty Water in KOH-Doped Hexagonal Ice: a Proton and Deuteron NMR Study. *Appl. Magn. Reson.* **2013**, *44*, 203−215.

https://dx.doi.org/10.1021/acs.jpcb.0c09266
J. Phys. Chem. B 2020, 124, 11491−11500