Rapid Communications

Dielectric spectrum of a DNA oligomer

Mithila V. Agnihotri[1] and Sherwin J. Singer[1,2]
[1]*Biophysics Program, Ohio State University, Columbus, Ohio 43210, USA*
[2]*Department of Chemistry and Biochemistry, Ohio State University, Columbus, Ohio 43210, USA*

We develop a hybrid method to extract the frequency-dependent dielectric susceptibility $\chi(\omega)$ from a 5-μs molecular dynamics trajectory of the Drew-Dickerson DNA dodecamer. Using nonparametric regression, polarization fluctuations are broken into slow and fast parts, each part treated by a different method optimal for its behavior. Our calculation reproduces the peak near 10^8 Hz in Im[$\chi(\omega)$] observed in recent experiments. This feature arises from the motion of the negatively charged DNA backbone and tightly bound counterions. Toward the low-frequency side of the peak, the DNA and counterions tend to move together and hence their contributions to the dielectric susceptibility cancel. At higher frequency the counterions disengage from the backbone and they each respond constructively to the time-varying electric field.

The structure and dynamical properties of DNA, intrinsic to its biological function, are reflected in its dielectric behavior. The static and frequency-dependent dielectric properties of DNA were studied even prior to the elucidation of its helical structure [1], when it was noted that the addition of DNA strongly increased the dielectric response of an aqueous solution. From that time until the present, the dielectric behavior of DNA has been intensively studied [2–4]. Besides fundamental scientific questions, dielectric properties of DNA are of interest for practical reasons. To give one example, the dielectric response provides a means to detect extremely low concentrations of DNA in micro- and nanofluidic devices [5–7].

Until recently, experiments were performed using DNA molecules that were too large to be accessible by detailed atomistic simulations. The situation changed in the past 11 years when dielectric spectra were obtained for 146-bp dsDNA oligomers [8], (10–120)-bp dsDNA [9], and (30−120)-bp ssDNA [10] (where bp denotes base pair, ds double-stranded, and ss single-stranded). The experimental spectra exhibit an absorption peak in the vicinity of 10^8 Hz for the shortest (10-bp) dsDNA oligomer [9]. The peak moves toward 10^7 Hz when the degree of dsDNA polymerization approaches the range of ∼100 base pairs [8,9]. Omori *et al.* interpreted this peak in terms of ionic motion relative to the DNA, considering both the strongly bound ions and distortion of the diffuse double layer [9]. For 146-bp dsDNA, Tomić *et al.* associate the response of the ion atmosphere with a much-lower-frequency feature near 8×10^3 Hz. Because of its sensitivity to DNA concentration, they assign the higher-frequency response in the $[(0.03–1.5) \times 10^7]$-Hz range to a collective structural property of a solution containing many chains.

From a 5-μs all-atom molecular dynamics trajectory, we calculate the dielectric response of the Drew-Dickerson dodecamer [d(CGCGAATTCGCG)], perhaps the most-well-studied DNA oligomer in x-ray, NMR and simulation studies. Our aim is to calculate the dielectric response at frequencies including the experimentally observed feature at 10^8 Hz, that is, on the 10-ns timescale, and offer insight into the origins of this feature. All simulations were performed using GROMACS 4.5.5 [11] in a periodically replicated simulation cell of dimension $(7.0042 \text{ nm})^3$, the volume achieved by first equilibrating the dodecamer (Protein Data Bank identifier 1BNA), 11257 SPC/E model [12] water molecules, 22 Na$^+$ ions required to neutralize the DNA, and six additional Na$^+$-Cl$^-$ pairs at a pressure of 1 bar. Over the run length of 5 μs, the temperature was maintained at 300 K using the Berendsen thermostat with $\tau = 3.33$ ps, the largest value of τ for which the temperature was stable.

Calculation of the dielectric response with expected frequencies in the 10^8-Hz range from atomistic simulations is a technically challenging task. Some difficulties arise because the dipole of an electrolyte in a periodically replicated simulation cell is not a well-defined quantity, i.e., it depends on the choice of simulation box for a periodic system containing ions [13,14]. Casting the dielectric susceptibility in terms of the current, which is independent of the choice of origin and simulation box position, avoids problems with the definition of dipole moment [15,16]. More serious difficulties arise because the statistical error estimating the time correlation functions required for the frequency-dependent dielectric susceptibility, say, the current-current correlation function $\langle J(t)J(0)\rangle$ from a trajectory of total length T,

$$\langle J(t)J(0)\rangle \approx \frac{1}{T-t} \int_0^{T-t} dt' \, J(t'+t)J(t'), \qquad (1)$$

rapidly increases with the time interval t. The root cause is that sample points $J(t'_1+t)J(t'_1)$ and $J(t'_2+t)J(t'_2)$ involve the same portion of the trajectory and are therefore correlated, unless $|t'_2 - t'_1| > t$. Therefore, the amount of statistically independent data available from a given trajectory for estimation of $\langle J(t)J(0)\rangle$ decreases as t increases. Still further difficulties derive from the long timescales associated with the motion

of large biomolecules. A previous attempt to characterize the frequency-dependent dielectric response of a DNA oligomer through atomistic simulation only used a brief trajectory of time length 7.3 ns [17].

Employing the current-current formulation because the dipole is not well defined in the presence of free charges, $\chi(\omega)$, the dielectric response tensor, is given as

$$\chi(\omega) = \frac{\beta}{\epsilon_0 V} \frac{i}{\omega} \int_0^\infty dt (e^{i\omega t} - 1) \langle \mathbf{J}(t) \mathbf{J}(0) \rangle, \qquad (2)$$

where $\mathbf{J} = \sum_{i=1}^N q_i \dot{\mathbf{r}}_i$, with q_i and $\mathbf{r}_i$ the charge and position of atom i, respectively. Another formulation of the dielectric response valid in the presence of free ions uses the dipolar displacement $\Delta \mathbf{M}(t_1, t_2)$ [13,16,18],

$$\Delta \mathbf{M}(t_1, t_2) = \mathbf{M}_I(t_2) - \mathbf{M}_I(t_1), \qquad (3)$$

where $\mathbf{M}_I(t) = \int_0^t dt' \mathbf{J}(t')$ is the itinerant dipole introduced by Caillol [13], the dipole of a system which begins in the primary simulation cell at $t = 0$ but whose particles are allowed to leak outside the cell instead of enforcing periodic boundary conditions. The dielectric susceptibility is given in terms of the dipolar displacement by the expression [16]

$$\chi(\omega) = \frac{\beta}{2\epsilon_0 V} \left\{ \mathbf{A} - i\omega \int_0^\infty dt\, e^{i\omega t} [\langle \Delta \mathbf{M}(0, t) \Delta \mathbf{M}(0, t) \rangle \right.$$
$$\left. - \mathbf{A} - \mathbf{B}t] \right\}, \qquad (4)$$

where

$$\lim_{t \to \infty} \langle \Delta \mathbf{M}(0, t) \Delta \mathbf{M}(0, t) \rangle = \mathbf{A} + \mathbf{B}t. \qquad (5)$$

Here $\langle \Delta \mathbf{M}(0, t) \Delta \mathbf{M}(0, t) \rangle$ is the mean-square dipolar displacement (MSDD). It is easily shown that the limiting linear form of the MSDD is related to the zero-frequency dielectric susceptibility and conductivity of the system

$$\chi(0) = \epsilon(0) - 1 = \frac{\beta}{2\epsilon_0 V} \mathbf{A}, \qquad (6)$$

$$\sigma(0) = \frac{\beta}{2} \mathbf{B}. \qquad (7)$$

Equation (7) links the conductivity to the long-time diffusive behavior of the MSDD. In the absence of free charges, $\mathbf{B} = 0$. The MSDD has all the advantages of an Einstein formulation of the susceptibility [16,18] and is strongly preferred over Eq. (2) because it requires far less data collected yet more effective sampling of a long trajectory [18]. Unfortunately, using the MSDD in Eq. (4) with free charges requires a fit to the long-time diffusive portion of the MSDD and, as noted by others [14] and has been our experience (see Supplemental Material [19]), is thus sensitive to the portion of the MSDD subject to large statistical uncertainty. To retain the advantages of the MSDD formulation yet avoid the instability of fitting its long-time diffusive behavior, we employed nonparametric regression [20] to break the itinerant dipole into slow and fast evolving parts

$$\mathbf{M}_{I,\text{slow}}(t) = \int_{-\infty}^\infty dt' g(t - t') \mathbf{M}_I(t'), \qquad (8)$$

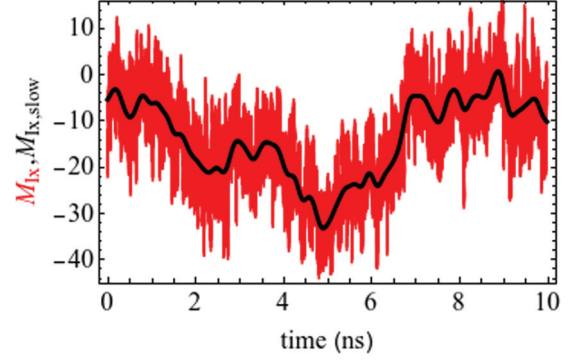$$\mathbf{M}_{I,\text{fast}}(t) = \mathbf{M}_I(t) - \mathbf{M}_{I,\text{slow}}(t), \qquad (9)$$

FIG. 1. The rapidly varying (red) curve is the x component of the itinerant dipole $\mathbf{M}_I(t)$. The slowly varying black curve is the x component of the smoothed $\mathbf{M}_{I,\text{slow}}(t)$ generated with a Gaussian smoothing kernel of $\sigma_g = 100$ ps. Here $\mathbf{M}_{I,\text{fast}}(t)$ is the difference between $\mathbf{M}_I(t)$ and $\mathbf{M}_{I,\text{slow}}(t)$.

where $g(t) = (2\pi\sigma_g^2)^{-1/2} \exp[-t^2/2\sigma_g^2]$ is a Gaussian smoothing kernel. The fast and slow parts are treated using different methods, each optimal for its respective behavior. As shown in Fig. 1, $\mathbf{M}_{I,\text{fast}}(t)$, because it fluctuates in a limited range relative to $\mathbf{M}_{I,\text{slow}}(t)$, is bounded, nondiffusive, and can be treated using Eq. (4) with $\mathbf{B} = 0$. In addition, $\mathbf{M}_{I,\text{slow}}(t)$ contains the diffusive behavior and is best treated by converting it into a current $\mathbf{J}_{\text{slow}}(t) = \int_{-\infty}^\infty dt' \dot{g}(t - t') \mathbf{M}_I(t')$ and then using $\mathbf{J}_{\text{slow}}(t)$ in Eq. (2), thereby avoiding the unstable fitting of the long-time MSDD. The susceptibility arising from $\mathbf{M}_{I,\text{fast}}(t)$ is the well-known $\sim 10^{11}$-Hz Debye peak of bulk water, and we do not consider it further. The slow-fast cross term is negligible for our choice of $\sigma_g = 100$ ps.

Before proceeding to results, a final methodological issue must be addressed. If the time integral in Eq. (2) is extended as far as allowed by the simulation length, the statistical noise in $\langle \mathbf{J}(t) \mathbf{J}(0) \rangle$ would eventually swamp any signal. Therefore, in practice, the integral must include a cutoff function $f(t)$, as in $\int_0^\infty dt\, f(t)(e^{i\omega t} - 1) \langle \mathbf{J}_{\text{slow}}(t) \mathbf{J}_{\text{slow}}(0) \rangle$.

Figure 2 shows the contribution from $\mathbf{J}_{\text{slow}}(t)$ to the dielectric susceptibility calculated using a Gaussian cutoff function $f(t) = \exp[-t^2/\sigma_f^2]$ with a sequence of σ_f values.
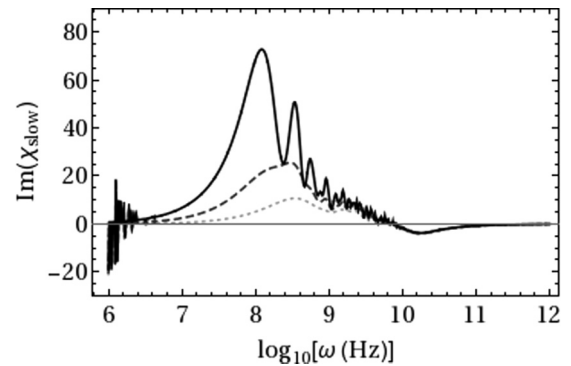

FIG. 2. Plot of Im$[\chi(\omega)]$ arising from $\mathbf{J}_{\text{slow}}(t)$ obtained using a Gaussian smoothing kernel with $\sigma_g = 100$ ps and Gaussian cutoff parameters $\sigma_f = 5$ ns (dotted line), $\sigma_f = 10$ ns (dashed line), and $\sigma_f = 20$ ns (solid line).
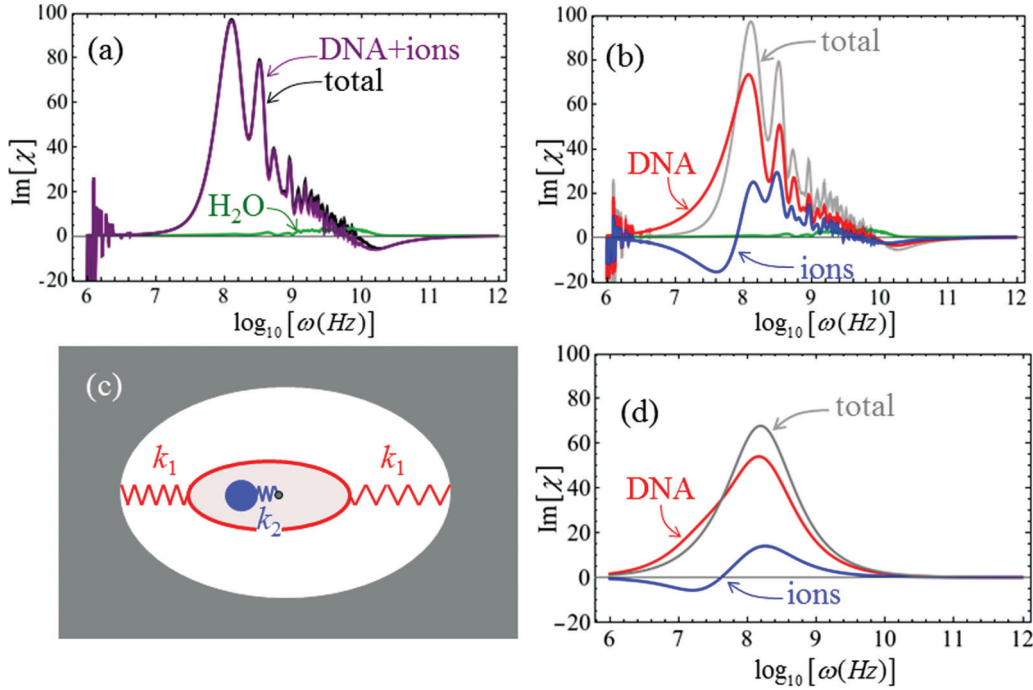
FIG. 3. Separation of $\mathrm{Im}[\chi_{\mathrm{slow}}(\omega)]$ into its components. (a) The total $\mathrm{Im}[\chi_{\mathrm{slow}}(\omega)]$ is separated into a water component and a DNA and ions component (b) The DNA and ions are further separated into a DNA backbone component and monatomic ions component. (c) Simple polarization model described in Eqs. (13) and (14) and (d) results of this model using the parameters given in the text. In (a) and (b) the slow component is generated using a Gaussian smoothing kernel with $\sigma_g = 100$ ps and $\sigma_f = 20$ ns for the Gaussian cutoff function.

Section S.3 of the Supplemental Material [19] demonstrates that smaller σ_f, i.e., earlier and more severe cutoffs, yield a smoother response function that is shifted to higher frequency. (Compare Fig. 2 with Fig. S4 of the Supplemental Material [19].) Longer cutoffs more faithfully reproduce the peak position but admit a sampling error which shows up as oscillations on the high-frequency side of the peak, exactly the behavior observed in Fig. 2.

Figure 2 confirms that simulations reproduce the feature near 10^8 Hz observed in experiments. For an interpretation, we separate $\mathbf{J}_{\mathrm{slow}}$ into components

$$\mathbf{J}_{\mathrm{slow}}(t) = \sum_{\mathrm{C = DNA,ions,H_2O}} \mathbf{J}_{\mathrm{slow,C}}(t), \qquad (10)$$

$$\mathbf{J}_{\mathrm{slow,C}} = \int_{-\infty}^{\infty} dt' \, \dot{g}(t - t') \mathbf{M}_{\mathrm{I,C}}(t'), \qquad (11)$$

where DNA refers to the DNA backbone and ions are all the monatomic ions, Na^+ and Cl^-. The contribution to the response function from each component is given by the correlation between that component's current and the total current [21]

$$\chi_{\mathrm{slow,C}}(\omega) = \frac{\beta}{\epsilon_0 V} \frac{i}{\omega} \int_0^{\infty} dt(e^{i\omega t} - 1)\langle \mathbf{J}_{\mathrm{slow,C}}(t)\mathbf{J}_{\mathrm{slow}}(0)\rangle f(t),$$
$$\mathrm{C = DNA,ions,H_2O}. \qquad (12)$$

Figure 3(a) separates $\mathrm{Im}[\chi_{\mathrm{slow}}(\omega)]$ (black line) into the response from all ionic species (purple line), that is, a DNA backbone plus monatomic ions, and the contribution of water (green line). Water makes a minor contribution at low frequency, and the response of all ionic species is not distinguishable from the total response. Projection of the slow part of the

current by smoothing is also a spatial filter for monatomic ions bound to DNA since those same species in bulk water have correlation times far less than 1 ps and frequency response near 10^{11} Hz [22].

In Fig. 3(b) we further separate the response of all ionic species into that of the DNA backbone (red line) and monatomic ions (blue line). At frequencies less than $\sim 10^8$ Hz, the response of these two groups cancels, while above this frequency they contribute constructively. This suggests that the DNA backbone and tightly bound counterions tends to move in phase and thus with opposing contributions to the response, at lower frequency, while at higher frequency they disengage and each contributes constructively to the response. How such behavior may arise is illustrated by a simple model with two degrees of freedom as shown in Fig. 3(c). One degree of freedom x_1, bearing charge q_1, represents motion of the DNA backbone. This degree of freedom moves relative to the surrounding solvent and monatomic ions far from the backbone [the outer gray area in Fig. 3(c)]. While the DNA motion is certainly diffusive at zero frequency, this degree of freedom acts as if restrained by a harmonic potential with force constant k_1 at frequencies of interest. The second degree of freedom x_2 represents monatomic ions closer to the DNA backbone, bearing charge q_2. This degree of freedom is restrained by a harmonic potential with force constant k_2 that depends on its position relative to the first degree of freedom. For simplicity in the following equations of motion, the friction coefficient ν for the two degrees of freedom is assumed equal:

$$\begin{pmatrix} m_1\ddot{x}_1 \\ m_2\ddot{x}_2 \end{pmatrix} = -\begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \nu\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix}. \qquad (13)$$

The coordinates x_1 and x_2 should not be taken literally as collinear displacements of DNA and ions along the DNA axis. Rather, the model merely illustrates how degrees of freedom can cross over from destructively to constructively contributing to the dielectric response as a function of frequency.

We expect that the system is in the overdamped limit, exhibiting exponential relaxation of coordinates to their equilibrium values, and set $\ddot{x}_1 = \ddot{x}_2 = 0$. In this limit, the model is easily solved using linear combinations of x_1 and x_2 that diagonalize the force constant matrix. We take $q_1 = -22e$, the DNA backbone charge, and anticipate that $|q_1| > |q_2|$ because only a fraction of the counterions are strongly bound to the DNA. The total charge of the background of ions far from DNA, represented by the gray area in Fig. 3(c), is $-(q_1 + q_2)$ with zero contribution to the dipole since it is symmetrically arranged about the origin. Hence, the total dipole moment of our model is $M(t) = M_{\mathrm{DNA}} + M_{\mathrm{ions}}$, where $M_{\mathrm{DNA}} = q_1 x_1(t)$ and $M_{\mathrm{ions}} = q_2 x_2(t)$. Using standard linear response theory, the imaginary part of the susceptibility is

$$\mathrm{Im}[\chi] = \frac{\beta\omega}{\epsilon_0 V} \int_0^\infty dt \cos(\omega t) \langle M_C(t) M(0) \rangle, \qquad (14)$$

where $C = \mathrm{DNA}$, ions, $M(t)$, its components evolve according to the overdamped limit of Eq. (13), and angular brackets indicate the Boltzmann average of the initial values of x_1 and x_2. Figure 3(d) shows typical results in which $q_2 = 11.84e$, $k_1 = 225.81\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$, $k_2 = 218.39\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$, and $\nu = 3661\,\mathrm{kJ\,mol^{-1}\,ns\,nm^{-2}}$. With these parameters, the lower-frequency in-phase eigenvector of the force constant matrix is $0.520x_1 + 0.854x_2$, while the higher-frequency out-of-phase eigenvector is $-0.854x_1 + 0.520x_2$. Figure 3(d) illustrates how in-phase motion leads to partial cancellation between DNA and monatomic ions at lower frequency and out-of-phase motion causes an additive response of DNA and ions at higher frequency.

In summary, we developed a hybrid method to extract the frequency-dependent dielectric susceptibility from a 5-μs molecular dynamics trajectory of the Drew-Dickerson DNA dodecamer. We avoided fitting the long-time diffusive behavior of the mean-square dipolar displacement while retaining the main advantage of the formalism, sparse sampling of data points. Our calculation reproduced the peak of $\mathrm{Im}[\chi(\omega)]$ near 10^8 Hz observed in recent experiments [9]. We were able to analyze the feature near 10^8 Hz and link it to motion of the negatively charged DNA backbone and tightly bound counterions. Toward the low-frequency side of the peak, the DNA and counterions tend to move together and hence their contributions to the dielectric susceptibility cancel. At higher frequency the counterions disengage from the backbone and they each respond constructively to the time-varying electric field.

Our interpretation of the 10^8-Hz feature is compatible with the assignment to ion motion by Omori *et al.* and provides a fuller understanding. Atomistic simulations are too limited by size, with respect to both the length of the oligomer and the size of the simulation box to evaluate the interpretation given by Tomić *et al.* To keep the number of solvent molecules within a feasible range, the effective concentration of our DNA oligomer is artificially large, corresponding to $35.4\,\mathrm{mg\,mL^{-1}}$. In the future, adaptive resolution methods should enable realistic simulation of the infinite dilution limit [23]. Omori *et al.* [9] examined concentration effects within a range of $1-4\,\mathrm{mg\,mL^{-1}}$ and found little qualitative change in the molar dielectric increment $\Delta\epsilon_m$ or the relaxation time τ, and concentration effects were smallest for the shortest oligomers. Tomić *et al.* [8] studied concentration effects for 146-bp dsDNA in pure water up to $5\,\mathrm{mg\,mL^{-1}}$ and in 1 mM NaCl up to $1.5\,\mathrm{mg\,mL^{-1}}$. Their spectra showed a marked concentration dependence, especially in their high-frequency feature at $\sim 10^7$ Hz, which presumably would correspond to the 10^8-Hz feature for shorter oligomers. Because of the concentration dependence, Tomić *et al.* assigned this feature to a collective structural process and assigned a lower-frequency peak to ion motion. Therefore, a definitive assessment of concentration effects requires extension of simulations to large system sizes without sacrificing accuracy.

[1] G. Jungner, I. Jungner, and L. G. Allgen, Nature (London) **163**, 849 (1949).

[2] M. Mandel and T. Odijk, Annu. Rev. Phys. Chem. **35**, 75 (1984).

[3] R. Hölzel, IET Nanobiotechnol. **3**, 28 (2009).

[4] M. Jiménez and T. Bellini, Curr. Opin. Colloid Interface Sci. **15**, 131 (2010).

[5] A. M. Johnson, D. R. Sadoway, M. J. Cima, and R. Langer, J. Electrochem. Soc. **152**, H6 (2005).

[6] S. M. Iqbal, D. Akin, and R. Bashir, Nat. Nanotechnol. **2**, 243 (2007).

[7] H. Ma, R. W. R. Wallbank, R. Chaji, J. Li, Y. Suzuki, C. Jiggins, and A. Nathan, Sci. Rep. **3**, 2730 (2013).

[8] S. Tomić, S. D. Babić, T. Ivek, T. Vuletić, S. Krča, F. Livolant, and R. Podgornik, Europhys. Lett. **81**, 68003 (2008).

[9] S. Omori, Y. Katsumoto, A. Yasuda, and K. Asami, Phys. Rev. E **73**, 050901(R) (2006).

[10] Y. Katsumoto, S. Omori, D. Yamamoto, A. Yasuda, and K. Asami, Phys. Rev. E **75**, 011911 (2007).

[11] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, Bioinformatics **29**, 845 (2013).

[12] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[13] J.-M. Caillol, J. Chem. Phys. **101**, 6080 (1994).

[14] M. Sega, S. S. Kantorovich, A. Arnold, and C. Holm, in *Recent Advances in Broadband Dielectric Spectroscopy*, edited by Y. P. Kalmykov, NATO Science for Peace and Security Series B: Physics and Biophysics (Springer, Dordrecht, 2012), pp. 103–122.

[15] J. M. Caillol, D. Levesque, and J. J. Weis, J. Chem. Phys. **85**, 6645 (1986).

[16] C. Schröder and O. Steinhauser, *Computational Spectroscopy* (Wiley-VCH, Weinheim, 2010), pp. 279–321.

[17] M. Ikeda, K. Nakazato, H. Mizuta, M. Green, D. Hasko, and H. Ahmed, Nanotechnology **14**, 123 (2003).

[18] M. V. Agnihotri, S.-H. Chen, C. Beck, and S. J. Singer, J. Phys. Chem. B **118**, 8170 (2014).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.98.060401 for analysis of numerical issues related to the extraction of the dielectric spectrum from a molecular dynamics trajectory.

[20] M. Wand and M. Jones, *Kernel Smoothing* (Chapman and Hall, New York, 1995).

[21] W. Bernard and H. B. Callen, Rev. Mod. Phys. **31**, 1017 (1959).

[22] S. Chowdhuri and A. Chandra, J. Chem. Phys. **115**, 3732 (2001).

[23] R. Podgornik, J. Zavadlav, and M. Praprotnik, Computation **6**, 3 (2018).

Correction: The title contained a misspelled word; the line following Eq. (3) contained a minor error. Both errors have been fixed.