# A multitechnique maximum entropy approach to the determination of the orientation and conformation of flexible molecules in solution

R. Berardi, F. Spinozzi, and C. Zannoni
*Dipartimento di Chimica Fisica e Inorganica, Università degli Studi di Bologna, Viale Risorgimento 4, 40136 Bologna, Italy*

(Received 9 January 1998; accepted 4 June 1998)

We present a maximum entropy method that allows the simultaneous analysis of different types of experimental data in order to obtain conformational information on flexible molecules in solution. We consider various NMR observables (dipolar, quadrupolar, J-couplings, nuclear Overhauser enhancements), and dielectric and neutron scattering techniques, and we express them using a common formalism in terms of orientational–conformational order parameters. We then show how these observables can be inverted in structural information allowing for continuous or discrete internal degrees of freedom and for any available prior information. We demonstrate the potentialities of the method on simulated $^1$HNMR, $^2$HNMR and dielectric data for some terminally halogenated alkyl chains and show the improvement in conformational analysis obtained by simultaneously analyzing different and complementary data sets. © *1998 American Institute of Physics.* [S0021-9606(98)50734-9]

## I. INTRODUCTION

The determination of the conformations of flexible molecules in solution and of their coupling to the overall molecular orientation is a classical problem in chemical physics that is receiving renewed experimental and theoretical attention.[1] On the experimental level, progress in the determination of homo-[2] and heteronuclear[3] dipolar couplings,[4,5] as well as of quadrupolar couplings[6] and nuclear Overhauser enhancements (NOE) data[7–10] has vastly improved the collection of available NMR tools.[11] Other classical techniques like the determination of quantities related to the dipole moment, such as the dielectric constant,[12,13] have also long served the quest for conformational determination. On the theoretical side, the extremely difficult task of extracting conformational information from the data has been tackled with a number of treatments.[6,14,15] However, most of these methods rely on the use of mean field theory and it is not always clear if and how this approximate treatment influences the results. In practice, the effective potential energy of a molecule in a liquid crystal solution $U(\omega,\Phi)$, depending on orientation $\omega$ and conformation $\Phi$, is written as

$$U(\omega,\Phi)=U_{\mathrm{int}}(\Phi)+U_{\mathrm{ext}}(\omega,\Phi), \tag{1}$$

where $U_{\mathrm{int}}(\Phi)$ is the intramolecular energy of the isolated molecule and $U_{\mathrm{ext}}(\omega,\Phi)$ represents the interaction between the molecule and the molecular field created by all the other molecules. In general, the internal part represents a property of the isolated molecule. For example, considering alkyl chains, as we are particularly interested in doing here, the first part could contain the internal energy difference $E_g$ between the *trans* and *gauche* states of a C–C bond in the gas phase. The external energy $U_{\mathrm{ext}}(\omega,\Phi)$ can in general be divided into an isotropic and an anisotropic term,

$$U_{\mathrm{ext}}(\omega,\Phi)=U_{\mathrm{ext}}^{\mathrm{iso}}(\Phi)+U_{\mathrm{ext}}^{\mathrm{aniso}}(\omega,\Phi). \tag{2}$$

The first term $U_{\mathrm{ext}}^{\mathrm{iso}}(\Phi)$ is often treated as a correction determined by the solvent to the intramolecular energy $U_{\mathrm{int}}(\Phi)$, while the second term contains the contribution to the total energy depending on the orientational order. In various approaches proposed in the literature,[1,2,6,16,17] a nonrigid molecule is considered as a collection of biaxial rigid conformers, whose ordering matrix is estimated using mean field theory. In the low solute concentration limit of this type of theory (see, e.g., Ref. 18), the effective orientational potential acting on a rigid molecule in a uniaxial phase is only a function of the orientational order parameters of the solvent molecules, normally assumed to be rigid and uniaxial and of a biaxiality parameter for the solute. The approaches that have been used in the literature differ in the way the biaxiality parameter is calculated for each conformer. Here we mention only a few of the most important ones: the *elastic continuum* model of Burnell and coworkers (VKB),[17] the *box shape* model of Straley,[19] the *surface tensor* model of Ferrarini *et al.*[15] and the *chord* model of Photinos *et al.*[6] In the elastic continuum model,[17] the molecular structure is approximated with a collection of van der Waals spheres placed at the atomic centers. The potential $U_{\mathrm{ext}}^{\mathrm{aniso}}(\omega,\Phi)$ represents the elastic energy, with the molecular deformation represented by the minimum circumference traced by the projection of the solute molecule in the conformer $\Phi$ onto a plane perpendicular to the director of nematic phase. This first *size and shape* model has been extended in various ways,[20,21] in particular taking into account the length of the projection of the molecule along the director. The model seems particularly suited to cases where the nematic solvent is one of the so-called *compensated mixtures* where the electric field gradient vanishes.[17,20,21] The Straley model[19] assumes that ordering is based on the molecular shape and approximates this as a rectangular box of a certain length (L), breadth (B) and width (W), calculated from the semiaxes

 

of the inertia ellipsoid, containing the molecule in the $\Phi$ conformation. The model also contains an adjustable parameter that characterizes the molecule–field interaction. The surface model[15] assumes that a local vector normal to each surface element of the solute tends to be aligned perpendicular to the director. The solute–solvent interaction tensor is then evaluated by estimating the surface area along different axes of the molecule. Finally we have the *chord* model for alkyl chains,[6] where the anisotropic part of the potential contains two terms that represent, respectively, the alignment of the separate C–C bond and the alignment of the *chord* connecting the midpoints of adjacent C–C bonds. These models have been employed to extract conformational information on various types of chains. In particular Pines and coworkers,[22] Polson and Burnell[21] and Luzar *et al.*[23] have used them to analyze NMR proton–proton dipolar couplings $D_{ij}$ data for various alkanes ranging from butane to decane in liquid crystal solution. An assumption central to these treatments is that the ordering of a conformer can be predicted by some single molecule property and its attendant molecular biaxiality. However, this is by no means established, even for a simple rigid molecule. For instance, we have recently examined the ability of various mean field models mentioned before to predict the ordering of simple rigid biaxial molecules, considering a set of 9,10 substituted anthracenes in the nematic ZLI1167, and found that none can actually reliably predict the observed biaxiality and its temperature dependence.[24] On a rather different note we have proposed[25] an approach that uses maximum entropy[26] to try to determine the flattest orientational–conformational distribution compatible with a set of experimental data. The method is a data inversion technique and does not rely on mean field theory and on the specific types of solute–solvent interactions being assumed. Maximum entropy methods have been applied by various authors to problems in dynamics,[27,28] as pioneered by Berne and to the determination of orientational distributions for rigid molecules.[29,30] We have applied the method to the analysis of NMR proton dipolar coupling data for simple rotameric molecules[31–33] and more recently to the much more complex case of flexible chains[34] in nematic solution. We have found that the approach can successfully recover conformer distributions even for alkyl chains provided certain favorable conditions exist. One is the availability of some prior information, for example the possibility of using the rotational isomeric state (RIS) model,[35] we have shown how to implement.[34] The other, and actually the only real limiting factor for the method, is the availability of a sufficiently large number of independent experimental results. NMR dipolar couplings constitute a particularly useful set of data, but in practice the complications associated with the analysis of the NMR spectra as the number of protons grows present serious limitations. Moreover, the NMR dipolar couplings do not contain purely scalar or rank other than second contributions that could provide additional terms needed for a full reconstruction of the conformational distribution [cf. Eq. (8)]. Another possibility, which is potentially the most promising, is that of performing various experiments using different complementary techniques and combining the results from the various approaches. For instance, deuterium NMR experiments provide average nuclear quadrupolar couplings that, even though mainly related to the C–D bond examined, can be affected by the orientation of the bond itself with respect to the molecular frame and thus, ultimately, by molecular conformations. For molecules possessing polar groups, the overall molecular dipole will vary with conformation, and indeed dipole moment analysis has been one of the first techniques used to attack the conformational problem.[12] Other potentially useful observables are $^3$J vicinal spin–spin couplings, nuclear Overhauser enhancements and neutron scattering (NSC) data. Clearly keeping into account the results coming from different techniques has often been attempted (see, e.g., Refs. 36 and 37). We believe, however, that the most powerful way of using various techniques efficiently is not that of analyzing each set of data independently in the usual way and only then combining the results. On the contrary, we think that combining the various input data and analyzing them together with the maximum entropy approach could provide the most effective way of obtaining an orientational–conformational distribution that complies to the available experimental knowledge. In this paper we wish to provide the basis for such a combined effort and we develop the required theoretical expressions using a common formalism for the various techniques. We shall also demonstrate the approach considering NMR dipolar couplings, NMR quadrupolar and electric dipole information in the analysis of some terminally halogenated alkyl chains.

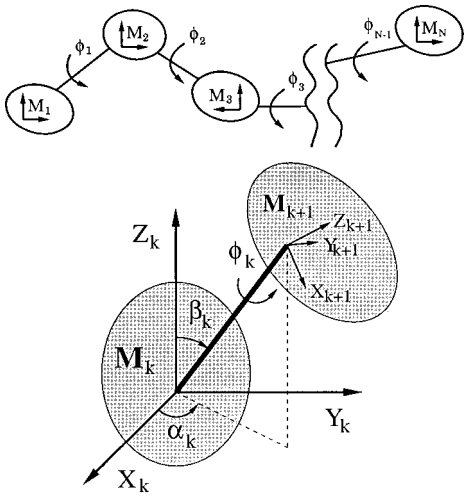


FIG. 1. Schematic representation of a general multirotor molecule showing the rigid frames $M_k$ and dihedral angles $\phi_k$ (top), and the spherical angles $\alpha_k$, $\beta_k$ (bottom) defining the orientation of bond $k$ linking the neighboring frames $M_k$ and $M_{k+1}$.

## II. THEORY

We consider the rather general case of a flexible multirotor molecule, treated as a set of $N$ rigid fragments linked by $N-1$ bonds[25,34] (cf. Fig. 1, top). Each rotor has its local reference system $M_k$ and the relative conformation of adjacent fragments $M_k$, $M_{k+1}$ is defined by a dihedral angle $\phi_k$. The overall conformational state is then specified by the set $\Phi \equiv (\phi_1, \ldots, \phi_{N-1})$, and we assume as reference conforma-

tion that with all angles $\phi_k=0$. In the case of *continuous* degrees of freedom, each angle $\phi_k$ can assume any real value in the range $[0,2\pi]$. However, in a number of cases, a *discrete* treatment, exemplified by the rotational isomeric state model of Flory,[35] in which the angles $\phi_k$ take only some (typically three) discrete values $\phi_k^{(j)}$ — or more simply $j_k$ — is well established. For alkyl chains, these values are the so-called *trans*, *gauche*$^+$ and *gauche*$^-$ states with $\phi_k=0$, $\pm 2\pi/3$ as typical values. The Euler angles $\omega\equiv(\alpha,\beta,\gamma)$[38] describe the rotation from the laboratory system LAB, with the $Z$ axis along the director of the mesophase, to the first (''*rigid*'') fragment $M_1$ and define accordingly the molecular orientation.

### A. Orientational–conformational distribution

The most complete information at a one–particle level on a flexible molecule in a uniform anisotropic solution is given by the singlet orientational–conformational distribution $P(\omega,\Phi)$, with the normalization condition

$$\int d\omega d\Phi P(\omega,\Phi)=1. \tag{3}$$

The purely conformational distribution $P(\Phi)$ is obtained by partial integration with respect to the orientational variables

$$P(\Phi)=\int d\omega P(\omega,\Phi). \tag{4}$$

The orientational–conformational distribution can be formally considered as an averaged product of Dirac delta functions that counts the particles in the various intervals

$$P(\omega,\Phi)=\langle \delta(\omega-\omega')\delta(\phi_1-\phi_1') \times \delta(\phi_2-\phi_2')\ldots\delta(\phi_{N-1}-\phi_{N-1}')\rangle_{\omega',\Phi'}, \tag{5}$$

where the symbol $\langle\ldots\rangle_{\omega',\Phi'}$ represents the average in the space of primed variables. Equation (5) can be written using the representation of the two types of angular $\delta$ functions in an orthogonal basis of Wigner matrices[38] as

$$\delta(\omega-\omega')=\sum_{L=0}^{\infty}\sum_{m=-L}^{L}\sum_{n=-L}^{L}\left(\frac{2L+1}{8\pi^2}\right)D^{L*}_{m,n}(\omega)D^{L}_{m,n}(\omega'), \tag{6}$$

and in a Fourier basis as

$$\delta(\phi_k-\phi_k')=\frac{1}{2\pi}\sum_{m=-\infty}^{\infty}e^{im(\phi_k-\phi_k')}. \tag{7}$$

In this way we obtain the expansion

$$P(\omega,\Phi)=\sum_{L,m,n;a_1,\ldots,a_{N-1}}\left[\frac{2L+1}{8\pi^2(2\pi)^{N-1}}\right] \times p^{L}_{m,n;a_1,\ldots,a_{N-1}}W^{L}_{m,n;a_1,\ldots,a_{N-1}}(\omega,\Phi), \tag{8}$$

where the mixed Wigner–Fourier basis functions $W$ are defined as

$$W^{L}_{m,n;a_1,\ldots,a_{N-1}}(\omega,\Phi) \equiv D^{L*}_{m,n}(\omega)\exp\{ia_1\phi_1+\ldots+ia_{N-1}\phi_{N-1}\}. \tag{9}$$

The expansion coefficients $p^{L}_{m,n;a_1,\ldots,a_{N-1}}$ are the *order parameters* for the orientational–conformational problem,

$$p^{L}_{m,n;a_1,\ldots,a_{N-1}}=\int d\omega d\Phi P(\omega,\Phi)W^{L*}_{m,n;a_1,\ldots,a_{N-1}}(\omega,\Phi) \tag{10}$$

$$\equiv\langle W^{L*}_{m,n;a_1,\ldots,a_{N-1}}(\omega,\Phi)_{\omega,\Phi}\rangle. \tag{11}$$

These order parameters form an infinite set that fully describes the molecular orientational–conformational order. In particular, we find as special cases the usual Saupe ordering matrix components $S_{zz}=\langle P_2\rangle=p^2_{0,0;0,\ldots,0}$, $S_{xx}-S_{yy}=\sqrt{6}\ \mathrm{Re}\langle D^2_{0,2}\rangle=\sqrt{6}\ \mathrm{Re}\ p^2_{0,2;0\ldots,0}$ and $S_{xz}=\sqrt{3/2}\ \mathrm{Re}\langle D^2_{0,1}\rangle=\sqrt{3/2}\ \mathrm{Re}\ p^2_{0,1;0,\ldots,0}$, where $S_{ij}$ are elements of the Saupe ordering matrix.[39] From the expansion coefficients we can compute the orientational order parameters measuring the average orientation of reference frame $M_k$ with respect to the LAB system. These are the order parameters of the single fragments $k$. For clarity, we consider simply connected structures and we call $\Omega_2$ the Euler angles measuring the orientation of frame $M_2$ with respect to $M_1$ in the reference conformation, $\Omega_3$ the orientation of frame $M_3$ with respect to $M_1$ and so on, until $\Omega_k$. We also call $\alpha_k$ and $\beta_k$ the polar angles, measured with respect to frame $M_k$, defining the local orientation of the chemical bond associated with the dihedral angle $\phi_k$, as shown in Fig. 1 (bottom). Using these definitions and measuring all orientations and angles with respect to the reference conformation we obtain the purely orientational order parameters of frame $M_k$ as

$$\langle D^{L}_{m,n}\rangle_{M_k}=\sum_{\substack{a_1,\ldots,a_{k-1}\\ b_1,\ldots,b_{k-1}}}p^{L}_{m,a_1;b_1,\ldots,b_{k-1},0,\ldots,0} \times G^{L}_{a_1,\ldots,a_{k-1},b_1,\ldots,b_{k-1},n}, \tag{12}$$

where we write $p^{L}_{m,a_1;b_1,\ldots,b_{k-1},0,\ldots,0}$ to indicate $p^{L}_{m,a_1;b_1,\ldots,b_{k-1},a_k,\ldots,a_{N-1}}$ with all subscripts, if any, from $k$ to $N-1$ equal to zero.

$$G^{L}_{a_1,\ldots,a_{k-1},b_1,\ldots,b_{k-1},n} =\sum_{r_1,\ldots,r_{k-1}}\left[\prod_{s=1}^{k-1}R^{L}_{a_s,b_s}(\alpha_s,\beta_s)R^{L*}_{r_s,b_s}(\alpha_s,\beta_s)\right] \times\left[\prod_{s=2}^{k-1}D^{L}_{r_{s-1},a_s}(\Omega_s)\right]D^{L}_{r_{k-1},n}(\Omega_k), \tag{13}$$

$$R^{L}_{m,n}(\alpha,\beta)\equiv e^{-in\alpha}d^{L}_{m,n}(\beta), \tag{14}$$

and $d^{L}_{m,n}(\beta)$ are small Wigner matrices.[38] The coefficients $G^{L}_{a_1,\ldots,a_{k-1},b_1,\ldots,b_{k-1},n}$ are defined in terms of geometrical parameters that are characteristic of the given molecular skeleton and are not modulated by molecular orientational and conformational changes. The orientational order parameters $\langle D^{L}_{m,n}\rangle_{M_k}$ are then a linear combination of coefficients $p^{L}_{m,n;a_1,\ldots,a_{N-1}}$, weighted by the constant geometrical $G$ co-

efficients. We can now apply these general transformation rules to the case of alkyl chains. For simplicity, we fix reference frame $M_1$ in a terminal molecular fragment. The remaining systems $M_k$ are collinear to $M_1$, with the $z_k$ axis parallel to the direction of full molecular elongation and the $x_k$ axis pointing along the symmetry axis of each $\widehat{\mathrm{HCH}}$ group, on the same side of the H atoms (cf. Fig. 2). Assuming the $\widehat{\mathrm{CCC}}$ angle $\theta$ to be equal for all fragments, we define $\psi \equiv (\pi-\theta)/2$, and using Eqs. (12)–(14), we write the $k$th frame order parameter as

$$\begin{aligned}\langle D^L_{m,n}\rangle_{M_k} = \sum_{\substack{a_1,\ldots,a_{k-1}\\ b_1,\ldots,b_{k-1}}} (-)^{a_2+a_4+\ldots+a_{[k/2]}} & \times p^L_{m,a_1;b_1,\ldots,b_{k-1},0,\ldots,0} \\ & \times d^L_{a_1,b_1}(\psi)\ldots d^L_{a_{k-1},b_{k-1}}(\psi) \\ & \times d^L_{a_2,b_1}(\psi)\ldots d^L_{n,b_{k-1}}(\psi), \end{aligned} \tag{15}$$

where $[k/2]$ is the integer part of $k/2$. In the case of *discrete* conformations, we introduce the singlet distribution function $P(\omega,\mathbf{j})$ representing the probability of finding a molecule with orientation within the range $[\omega,\omega+d\omega]$ and conformation $\mathbf{j}\equiv(j_1,\ldots,j_{N-1})$. In this case the normalization condition becomes

$$\sum_{\mathbf{j}} \int d\omega P(\omega,\mathbf{j}) = 1. \tag{16}$$

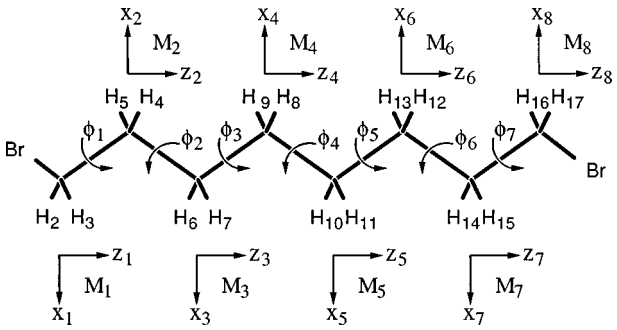


FIG. 2. The rigid fragment coordinate frames $M_k$, the dihedral angles $\phi_k$ and the proton numbering scheme adopted for the terminal di-bromo–*n*–alkanes (1,8-di–bromo–octane shown). The reference conformation is all–*trans*.

Again, the conformational distribution $P(\mathbf{j})$ is obtained by partial integration

$$P(\mathbf{j}) = \int d\omega P(\omega,\mathbf{j}). \tag{17}$$

The orientational–conformational order parameters are

$$p^L_{m,n;j_1,\ldots,j_{N-1}} = \int d\omega P(\omega,\mathbf{j}) D^L_{m,n}(\omega), \tag{18}$$

and they are the expansion coefficients of the orientational distribution function for the molecule in a fixed conformation $\mathbf{j}$. For alkyl chains described using the *discrete* RIS model [cf. Eq. (12)], we have

$$\begin{aligned}\langle D^L_{m,n}\rangle_{M_k} = \sum_{\mathbf{j}} p^L_{m,n;j_1,\ldots,j_{N-1}} \sum_{\substack{a_1,\ldots,a_{k-1}\\ b_1,\ldots,b_{k-1}}} (-)^{a_2+a_4+\ldots+a_{[k/2]}} & e^{\{-ib_1\phi_1^{(j_1)}-\ldots-ib_{k-1}\phi_{k-1}^{(j_{k-1})}\}} \\ & \times d^L_{a_1,b_1}(\psi)\ldots d^L_{a_{k-1},b_{k-1}}(\psi) d^L_{a_2,b_1}(\psi)\ldots d^L_{n,b_{k-1}}(\psi). \end{aligned} \tag{19}$$

The knowledge of these order parameters $\langle D^L_{m,n}\rangle_{M_k}$ is sufficient to calculate any single particle and bond observable of an alkyl chain. For instance, the order parameter for a CH bond in a methylene group, often measured from $^2$HNMR experiment as $S_{\mathrm{CD}}$, can be written as $S_{\mathrm{CD}} = -\frac{1}{2}S_{zz} + \frac{1}{2}(S_{xx}-S_{yy})\cos\chi$, where $S_{zz} = \langle D^2_{0,0}\rangle_{M_k}$, $S_{xx}-S_{yy} = \sqrt{6}\ \mathrm{Re}\langle D^2_{0,2}\rangle_{M_k}$ and $\chi$ is the $\widehat{\mathrm{HCH}}$ bond angle.

### B. Multitechnique data combination

We now introduce a rather general formalism for the simultaneous analysis of experimental data sets measured with different techniques. To do so, it is necessary to devise a suitable symmetrization and to determine a scheme of multitechnique linearly independent combination of observables. The symbol $\mathcal{F}_{\mu,i}$ identifies the $i$–th observable in the data set provided by technique $\mu$, and described by the function $[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)$ referred to a common laboratory frame. The experimental measure of each observable is related to the average of the corresponding function over the unknown single particle orientational–conformational distribution $P(\omega,\Phi)$,

$$\langle [\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}\rangle = \int d\omega d\Phi [\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi) P(\omega,\Phi). \tag{20}$$

The functions $\mathcal{F}_{\mu,i}$ can be expanded in terms of irreducible spherical components as

$$[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi) = \sum_{L=0}^{\infty} \sum_{m=-L}^{L} [\mathcal{F}_{\mu,i}]^{L,m}_{\mathrm{LAB}}(\omega,\Phi), \tag{21}$$

where the number of tensorial components actually contributing to the expansion is dictated by the particular observable $[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}$. For instance, considering second rank NMR observables, only the components with $L=2$ can be different from zero. On the other hand, for measurements coming from scattering techniques (e.g., x–rays or neutrons), com-

ponents of every rank $L$ can, in principle, contribute. Transforming spherical components to frame $M_1$, it is possible to write an equivalent expansion,

$$[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)=\sum_{L=0}^{\infty}\sum_{m=-L}^{L}\sum_{n=-L}^{L} D^{L*}_{m,n}(\omega) \times[\mathcal{F}_{\mu,i}]^{L,n}_{M_1}(\Phi). \tag{22}$$

All observable functions $[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)$ should be symmetrized according to both molecular symmetry and the other transformations characteristic of the measurement. This operation could reduce the number of observable data. For instance, in the case of a molecule with two identical rotors, one needs to symmetrize the NMR dipolar or quadrupolar couplings with respect to a fragment exchange.[33] There are at least two equivalent ways of performing the symmetrization. The first, more general method, applies projection operators to accomplish symmetrization with respect to the operations corresponding to the various degrees of freedom, that is: the local conformational symmetry of the single rotor[40,41]; the symmetry of the whole molecule in an arbitrary conformational state $\Phi$; the symmetry of the mesophase and, finally, that of the experiment. The second method is often used, for instance, in NMR spectra interpretation. It is based on equivalent nuclei or pairs of nuclei and implies averaging over these. The two schemes give completely equivalent results. An example of this second symmetrization scheme, using NMR techniques, will be given in Secs. III A–III C. In order to perform a combined analysis using functions representing different physical properties (i.e., with different technique labels $\mu$), it is useful to transform these functions to a dimensionless form. To do this, for each technique $\mu$ we find the maximum absolute value — $\max\{|[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)|\}_{\mu}$ — of all functions $i$ over a certain orientational–conformational grid and we use this result to renormalize each observable $\mathcal{F}_{\mu,i}$. Using this prescription all functions are scaled within the dimensionless range $[-1,1]$. In addition, using observables with similar magnitudes is preferable from a purely numerical point of view. The combined scaling–symmetrization procedure leads to

$$[\mathcal{F}^{S}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)=\frac{w_\mu}{\max\{|[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi)|\}_{\mu}} \times\frac{1}{n_S}\sum_{S=1}^{n_S}\mathcal{O}_S[\mathcal{F}_{\mu,i}]_{\mathrm{LAB}}(\omega,\Phi), \tag{23}$$

where each projection operator $\mathcal{O}_S$ corresponds to one of the $n_S$ symmetry operations, and $w_\mu$ is a factor that can in principle be used to weight the different techniques $\mu$ according to their sensitivity and reliability. Here we shall use $w_\mu=1$. We suppose now to have $N_\mu$ experimental data for the $\mu$–th technique. The total number of symmetrized observables becomes $N_{\mathcal{F}}=\Sigma_{\mu=1}^{M}N_\mu$, where $M$ is the number of different techniques considered. The orientational–conformational information that can be extracted from the whole group of different techniques depends on $N_{\mathcal{F}}$ but not all of these observables will necessarily add new information. Thus, before considering the different experimental data and analyzing them, it is important to determine a set of suitable linearly independent combinations. A systematic way of doing this is through the introduction of a scalar product between two fixed observable functions in the LAB frame.[33] For the *continuous* case we have

$$\begin{aligned}([\mathcal{F}^{S}_{\mu,i}]_{\mathrm{LAB}}|[\mathcal{F}^{S}_{\mu',i'}]_{\mathrm{LAB}}) &= \sum_{L,m,L',m'}\int d\omega d\Phi[\mathcal{F}^{S*}_{\mu,i}]^{L,m}_{\mathrm{LAB}}(\omega,\Phi) \times[\mathcal{F}^{S}_{\mu',i'}]^{L',m'}_{\mathrm{LAB}}(\omega,\Phi) \\ &= \sum_{L,n}\frac{8\pi^2}{2L+1}\int d\Phi[\mathcal{F}^{S*}_{\mu,i}]^{L,n}_{M_1}(\Phi)[\mathcal{F}^{S}_{\mu',i'}]^{L,n}_{M_1}(\Phi),\end{aligned} \tag{24}$$

where the simplification has been carried out using the orthogonality of Wigner matrices.[38] The integration with respect to the conformational variables can be computed numerically. For the *discrete* conformations case we obtain instead

$$\begin{aligned}([\mathcal{F}^{S}_{\mu,i}]_{\mathrm{LAB}}|[\mathcal{F}^{S}_{\mu',i'}]_{\mathrm{LAB}}) &= \sum_{L,m,L',m'}\int d\omega\sum_{\mathbf{j}}[\mathcal{F}^{S*}_{\mu,i}]^{L,m}_{\mathrm{LAB}}(\omega,\mathbf{j})[\mathcal{F}^{S}_{\mu',i'}]^{L',m'}_{\mathrm{LAB}}(\omega,\mathbf{j}) \\ &= \sum_{L,n}\frac{8\pi^2}{2L+1}\sum_{\mathbf{j}}[\mathcal{F}^{S*}_{\mu,i}]^{L,n}_{M_1}(\mathbf{j})[\mathcal{F}^{S}_{\mu',i'}]^{L,n}_{M_1}(\mathbf{j}).\end{aligned} \tag{25}$$

The scalar products $V_{IJ}=([\mathcal{F}^{S}_{I}]_{\mathrm{LAB}}|[\mathcal{F}^{S}_{J}]_{\mathrm{LAB}})$ are the elements of a symmetric $N_{\mathcal{F}}\times N_{\mathcal{F}}$ overlap matrix $\mathbf{V}$, where we have introduced the symbol $I$ and $J$ to label the pair $(\mu,i)$ and $(\mu',i')$. The dimension $N_{\mathcal{G}}$ (with $N_{\mathcal{G}}\leqslant N_{\mathcal{F}}$) of the function space spanned by the observables is then found diagonalizing $\mathbf{V}$ using standard techniques of linear algebra. The $N_{\mathcal{K}}$ eigenvalues which are zero within a given threshold $\tau\geqslant 0$ are discarded and the remaining $N_{\mathcal{G}}=N_{\mathcal{F}}-N_{\mathcal{K}}$ orthogonal eigenvectors corresponding to the eigenvalues larger than $\tau$ are normalized. We call $\mathbf{Z}$ the $N_{\mathcal{F}}\times N_{\mathcal{G}}$ matrix containing the $N_{\mathcal{G}}$ eigenvectors as its columns. Thus we identify a set of $N_{\mathcal{G}}$ orthogonal basis functions,

$$\begin{aligned}[\mathcal{G}_I]_{\mathrm{LAB}}(\omega,\Phi) &= \sum_{J=1}^{N_{\mathcal{F}}}[\mathcal{F}^{S}_{J}]_{\mathrm{LAB}}(\omega,\Phi)Z_{JI} \\ &= \sum_{L=0}^{\infty}\sum_{m=-L}^{L}\sum_{n=-L}^{L}D^{L*}_{mn}(\omega)[\mathcal{G}_I]^{L,n}_{M_1}(\Phi).\end{aligned} \tag{26}$$

In a similar fashion this transformation rule defines the linear combinations $\langle[\mathcal{G}_I]_{\mathrm{LAB}}\rangle$ of the experimental data to be used as observables,

$$\langle[\mathcal{G}_I]_{\mathrm{LAB}}\rangle=\sum_{J=1}^{N_{\mathcal{F}}}\langle[\mathcal{F}^{S}_{J}]_{\mathrm{LAB}}\rangle Z_{JI}. \tag{27}$$

### C. Combined techniques maximum entropy distribution

Following information theory[26] we can define the entropy functional associated with a certain distribution as

$$S[P(\omega,\Phi)]=-\int d\omega d\Phi P(\omega,\Phi)\log[P(\omega,\Phi)]. \quad (28)$$

According to maximum entropy[42,43] (ME), the best (least biased) approximation to the true distribution that can be inferred from the experimental data is

$$P(\omega,\Phi)=\frac{1}{Z_0}\exp\left\{\sum_{I=1}^{N_{\mathcal{G}}}\lambda_I[\mathcal{G}_I]_{\mathrm{LAB}}(\omega,\Phi)\right\}, \quad (29)$$

where we have used the functions $[\mathcal{G}_I]_{\mathrm{LAB}}(\omega,\Phi)$ corresponding to the linearly independent experimental observables $\langle[\mathcal{F}_I^S]_{\mathrm{LAB}}\rangle$. Notice that the set of available experiments determines the number $N_{\mathcal{F}}$ of observables $\langle[\mathcal{F}_I^S]_{\mathrm{LAB}}\rangle$ and their $N_{\mathcal{G}}$ linearly independent combinations $\langle[\mathcal{G}_I]_{\mathrm{LAB}}\rangle$. The normalization term $Z_0$ is defined as

$$Z_0(\{\lambda\})=\int d\omega d\Phi\ \exp\left\{\sum_{I=1}^{N_{\mathcal{G}}}\lambda_I[\mathcal{G}_I]_{\mathrm{LAB}}(\omega,\Phi)\right\}, \quad (30)$$

and $\{\lambda\}\equiv\{\lambda_1,\ldots,\lambda_{N_{\mathcal{G}}}\}$ is a set of variational parameters. The practical determination of the maximum entropy distribution is performed defining a suitable free energylike convex functional,[42,43]

$$\Gamma(\{\lambda\})=\ln Z_0(\{\lambda\})-\sum_{I=1}^{N_{\mathcal{G}}}\lambda_I\langle[\mathcal{G}_I]_{\mathrm{LAB}}\rangle, \quad (31)$$

and optimizing the set of $\{\lambda\}$ parameters until the absolute minimum of $\Gamma(\{\lambda\})$ is found. The convexity of the pseudo-free energy $\Gamma(\{\lambda\})$ ensures the existence of the optimal solution.[43] Equation (29) is the maximum entropy distribution function when no *a priori* knowledge for the system is available. The maximum entropy approach does not require further assumptions in the case that a sufficiently high number of data exists. On the other hand, since the maximum entropy method strictly performs an inversion of experimental data it cannot be of help if no such information exists. As a consequence of its nature, the maximum entropy distribution is inevitably isotropic (i.e., flat) when no experimental data (i.e., $N_{\mathcal{G}}=0$) are available. However, as is often the case, some *intrinsic* information, or well assessed knowledge from previous investigations, may be available and the experiments at hand could actually be used to complement this information rather than being required to ignore it.[34] This *prior* information could be the knowledge of some important details on the molecular structure, e.g., van der Waals radii or other geometrical constraints hindering certain conformers. Constraints obtained from molecular dynamics simulations have also been recently proposed.[44] If the intramolecular energy of the isolated molecule $U_{\mathrm{int}}(\Phi)$ is known we could even calculate a complete *intrinsic*[45] conformational probability $P_i(\Phi)$. In general, following a Bayesian approach,[45] we can then write

$$P(\omega,\Phi)=(1/Z)P_i(\Phi)P_e(\omega,\Phi|P_i), \quad (32)$$

where the *a priori* distribution $P_i(\Phi)$ plays the role of a weight factor for the conformational states, resulting from the known intramolecular interactions for the molecular skeleton. The second term $P_e$ is the *external* (i.e., *posterior*) distribution that we estimate applying the maximum entropy principle to the analysis of experimental data, while the constant $Z$ allows the distributions to be separately normalized. In condensed fluid phases, $P_e(\omega,\Phi|P_i)$ describes how packing,[46] and more generally the solvent, affects the actual configuration $(\omega,\Phi)$ given the underlying distribution $P_i(\Phi)$ for the isolated molecule. The application of the maximum entropy algorithm is then restricted to the part of the full orientational–conformational distribution which is still unknown, namely $P_e$, described by Eq. (29). Furthermore, the *intrinsic* distribution plays the role of a weight factor for conformations, and the overlap matrix elements [cf. Eq. (24)] become

$$V_{IJ}=\sum_{L,n}\frac{8\pi^2}{2L+1}\int d\Phi P_i(\Phi)[\mathcal{F}_I^{S*}]_{M_1}^{L,n}(\Phi)[\mathcal{F}_J^S]_{M_1}^{L,n}(\Phi). \quad (33)$$

All integrals over the conformational variables $\Phi$ are modified in a similar fashion to include the *intrinsic* distribution $P_i(\Phi)$. Using a *discrete* RIS approximation the orientational–conformational distribution [cf. Eq. (32)] becomes

$$P(\omega,\mathbf{j})=(1/Z)P_i(\mathbf{j})P_e(\omega,\mathbf{j}|P_i). \quad (34)$$

Again, the three states *gauche*$^{\pm}$ and *trans* can be assumed to have the same *a priori* probability. More generally, taking advantage of previous knowledge allows writing the *intrinsic* distribution as

$$P_i(\mathbf{j})\propto f(\mathbf{j})\exp\left\{\sum_{k=1}^{N-1}E_{g_k}^*\delta_{j_k,t}\right\}, \quad (35)$$

where $E_{g_k}^*=E_{g_k}/k_BT$ is the dimensionless energy of a *gauche* with respect to *trans* for the $k$th conformational bond at temperature $T$ ($k_B$ is the Boltzmann constant) and $\delta_{j_k,t}=1$ if the state is *trans* ($t$) and zero otherwise. Other *a priori* conformational knowledge can be easily built in by means of the filter function $f(\mathbf{j})$. For instance, this function can be taken to be zero or one if a conformation is sterically hindered or not, leading to the number $n_T$ of accessible conformations ($n_T\leqslant 3^{N-1}$). The maximum entropy analysis can also be performed taking into account the extent of experimental data uncertainties and their effect on the distribution function compatible with the available physical observables. To do so, we assume that all $N_{\mathcal{F}}$ experimental observable functions $\langle[\mathcal{F}_I^S]_{\mathrm{LAB}}^{L,m}\rangle$ have been sampled from Gaussian distributions of variance $\sigma_i$ (estimated from experimental error ranges). We then generate $M$ additional data sets by sampling with a Monte Carlo method from $N_{\mathcal{F}}$ Gaussians of width $\sigma_i$ centered at the experimental values $\langle[\mathcal{F}_I^S]_{\mathrm{LAB}}^{L,m}\rangle$. The sampling intervals are generally $\pm p\sigma_i$, with $p=3$, while $M=50$ is usually sufficient for most cases.[33] Each generated data set is then separately analyzed with the maximum entropy algorithm to obtain a distribution $P^{(m)}(\omega,\Phi)$. The resulting set of $M+1$ distributions is finally combined to define the average orientational–conformational distribution

$$\overline{P}(\omega,\Phi)=\frac{1}{M+1}\sum_{m=0}^{M}P^{(m)}(\omega,\Phi). \quad (36)$$

Using this prescription it is possible to compute other distribution functions for the system studied. For instance, from each $P^{(m)}(\omega,\Phi)$ we can obtain by partial integration the distribution $P^{(m)}(\phi_k)$ for a single conformational angle $\phi_k$, and after averaging, the distribution $\overline{P}(\phi_k)$. We can also calculate for each configuration $(\omega,\Phi)$ a local standard deviation using the equation

$$\sigma^2_{\omega,\Phi}=\frac{1}{M}\sum_{m=0}^{M}[P^{(m)}(\omega,\Phi)-\overline{P}(\omega,\Phi)]^2, \tag{37}$$

and thus we can estimate the uncertainty on our maximum entropy results. The same procedure can be employed to calculate the error on order parameters.

## III. PHYSICAL OBSERVABLES

The combined approach can be used, in principle, with any experimental technique, providing physical observables modulated by the conformational distribution. Here, we consider various techniques frequently used independently and we briefly describe how their observables can be treated in a unified way with a maximum entropy analysis.

### A. NMR quadrupolar splittings

The first observables we examine are the nuclear quadrupolar splittings, measured from NMR spectra in anisotropic solvents.[4,6,16,37,47–49] A nucleus $k$ with spin $I_k \geqslant 1$, notably a deuteron $^2$H, has an electric quadrupole moment different from zero. For a nuclear spin $I_k$ in a given local chemical environment (e.g., alyphatic or aromatic C–D bond, etc.) the quadrupole coupling tensor $\mathbf{q}_k$ is[11]

$$\mathbf{q}_k=\frac{eQ_k}{2\hbar I_k(2I_k-1)}\mathbf{V}_k, \tag{38}$$

where $Q_k$ and $\mathbf{V}_k$ are the quadrupole moment and the electric field gradient tensor at the site of nucleus $k$, measured with respect to the molecular reference frame $M_q$, which makes $\mathbf{q}_k$ diagonal and with biaxiality $\eta_k \equiv ([q_k]_{yy}-[q_k]_{xx})/[q_k]_{zz}$. The tensor $\mathbf{q}_k$ is traceless and thus it can be measured only in anisotropic phases. In particular, for uniaxial liquid crystals, the only relevant component is $\langle[q_k]_{ZZ}\rangle$, where $Z$ is the laboratory magnetic field direction that we take to be parallel to the mesophase director. Using irreducible spherical components we can write

$$[q_k]^{2,0}_{\mathrm{LAB}}(\omega,\Phi)=[q_k]^{2,0}_{M_q}\sum_{n=-2}^{2}D^{2*}_{0,n}(\omega)\left\{D^{2*}_{n,0}[\omega_q(\Phi)]-\frac{\eta_k}{\sqrt{6}}[D^{2*}_{n,2}[\omega_q(\Phi)]+D^{2*}_{n,-2}[\omega_q(\Phi)]]\right\}, \tag{39}$$

where $\omega_q$ represents the Euler angles that describe the rotation from the system $M_1$ to the system $M_q$. For a single deuteron the NMR spectrum contains a doublet, and the average value of this quadrupolar component obtained from the splitting can be written as

$$\langle[q_k]^{2,0}_{\mathrm{LAB}}\rangle=[q_k]^{2,0}_{M_q}\sum_{n=-2}^{2}\left\{D^{2*}_{n,0}(\omega'_q)-\frac{\eta_k}{\sqrt{6}}[D^{2*}_{n,2}(\omega'_q)+D^{2*}_{n,-2}(\omega'_q)]\right\}\times\sum_{\substack{a_1,\ldots,a_{j-1}\\ b_1,\ldots,b_{j-1}=2}}^{2}p^{2*}_{0,a_1;b_1,\ldots,b_{j-1},0,\ldots,0}\times G^{2*}_{a_1,\ldots,a_{j-1},b_1,\ldots,b_{j-1},n}, \tag{40}$$

where now $\omega'_q$ are the Euler angles that describe the rotation from the system $M_j$ (in which we suppose the spin $k$) to $M_q$. We see that only a finite subset of order orientational–conformational parameters is sufficient to fully describe the average quadrupolar splittings. The quadrupolar tensor biaxiality $\eta_k$ is usually small and it is common practice, at least for aliphatic chains, to use $\eta_k=0$ in $^2$HNMR calculations. The quantities directly measured from the spectrum are the quadrupolar splittings $\Delta\nu_k$ (i.e., the separation of a deuteron doublet) related to the averaged spherical components as

$$\Delta\nu_k=\frac{3}{2}\langle[q_k]_{ZZ}\rangle\equiv\left(\frac{3}{2}\right)^{1/2}\langle[q_k]^{2,0}_{\mathrm{LAB}}\rangle. \tag{41}$$

If the bond orientation $\omega_q$ is modulated by some conformational degrees of freedom, the tensor $\mathbf{q}_k$ contains information on both molecular orientation $\omega$ and conformational structure $\Phi$. If frames $M_1$ and $M_q$ are equivalent, we have

$$\Delta\nu_k=\frac{3eV_kQ_k}{4\hbar I_k(2I_k-1)}\langle P_2(\cos\,\theta_k)\rangle, \tag{42}$$

where $\theta_k$ is the angle between the C–D bond and magnetic field directions, and $V_k$ is the $\mathbf{z}$ component of $\mathbf{V}_k$ measured in frame $M_q$. As already mentioned in Sec. II B, the symmetrization of noninteracting quadrupolar coupling $[q_k]^{2,0}_{\mathrm{LAB}}$ is usually performed considering the set of $n_P$ *equivalent* deuterons. In this case, the symmetrized dimensionless quadrupolar coupling becomes

$$[q^S_k]^{2,0}_{\mathrm{LAB}}(\omega,\Phi)=\left(\frac{2}{3}\right)^{1/2}\frac{2\hbar I_k(2I_k-1)}{eV_kQ_k}\times\frac{1}{n_P}\sum_{P=1}^{n_P}[q_{\{k\}_P}]^{2,0}_{\mathrm{LAB}}(\omega,\Phi), \tag{43}$$

where $\{k\}_P$ is one of the $n_P$ equivalent deuterons.

### B. NMR dipolar couplings

The dipolar coupling tensor $\mathbf{T}_{ij}$ between a pair of nuclei $i$ and $j$ with gyromagnetic ratios $\gamma_i$ and $\gamma_j$ can be written as[4]

$$\mathbf{T}_{ij}=-\frac{h\gamma_i\gamma_j}{8\pi^2r^5_{ij}}(3\mathbf{r}_{ij}\otimes\mathbf{r}_{ij}-r^2_{ij}\mathbf{I}), \tag{44}$$

where $h$ is the Planck constant, $\mathbf{I}$ is the $3\times3$ identity matrix, and $\mathbf{r}_{ij}=\mathbf{r}_j-\mathbf{r}_i$ is the internuclear vector of length $r_{ij}$ defining the relative position of nuclei $i$ and $j$ with respect to the laboratory reference frame. Since the trace of $\mathbf{T}_{ij}$ is zero, NMR dipolar effects cannot be measured in isotropic fluid

phases. Anisotropic uniaxial solvents are commonly used to record NMR spectra of probe molecules, and the averaged $\langle [T_{ij}]_{ZZ} \rangle$ are the only nonzero components measurable in these uniaxial phases, where $Z$ is the laboratory magnetic field direction, which is supposed to be collinear to the mesophase director. These averages are related to the dipolar couplings $D_{ij}$ as

$$D_{ij} \equiv \langle [T_{ij}]_{ZZ} \rangle = (\tfrac{2}{3})^{1/2} \langle [T_{ij}]^{2,0}_{\mathrm{LAB}} \rangle = -\frac{h\gamma_i\gamma_j}{4\pi^2} \left\langle \frac{P_2(\cos\theta_{ij})}{r_{ij}^3} \right\rangle, \tag{45}$$

where $P_2$ is the second Legendre polynomial and $\theta_{ij}$ is the angle between the internuclear vector $\mathbf{r}_{ij}$ and the direction of the magnetic field. In general, both $\mathbf{r}_{ij}$ and $\theta_{ij}$ will depend on molecular orientation and conformation, thus dipolar couplings can be useful for recovering geometrical distribution functions. The spherical components $[T_{ij}]^{2,0}_{\mathrm{LAB}}(\omega,\Phi)$ measured in the laboratory frame are related to those referred to the first molecular system $M_1$,

$$[T_{ij}]^{2,0}_{\mathrm{LAB}}(\omega,\Phi) = \sum_{n=-2}^{2} D^{2*}_{0,n}(\omega)[T_{ij}]^{2,n}_{M_1}(\Phi) = -\left(\frac{3}{2}\right)^{1/2} \frac{h\gamma_i\gamma_j}{4\pi^2} \sum_{n=-2}^{2} D^{2*}_{0,n}(\omega) \times \frac{D^{2*}_{n,0}[\omega_{ij}(\Phi)]}{r_{ij}^3(\Phi)}. \tag{46}$$

The factor $[T_{ij}]^{2,n}_{M_1}(\Phi)$ could be expanded in Fourier series to give

$$[T_{ij}]^{2,n}_{M_1}(\Phi) = \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} [T_{ij}]^{2,n}_{a_1,\ldots,a_{N-1}} \times \exp\{ia_1\phi_1 + \ldots + ia_{N-1}\phi_{N-1}\}, \tag{47}$$

where the coefficients $[T_{ij}]^{2,n}_{a_1,\ldots,a_{N-1}}$ are fixed once the skeleton form is given. Substituting in Eq. (46) and averaging, we obtain the average dipolar couplings in terms of an infinite set of order parameters

$$\langle [T_{ij}]^{2,0}_{\mathrm{LAB}} \rangle = \sum_{n=-2}^{2} \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} p^{2*}_{0,n;a_1,\ldots,a_{N-1}} \times [T_{ij}]^{2,n}_{a_1,\ldots,a_{N-1}}. \tag{48}$$

Similarly to Eq. (43), the spherical components of dipolar couplings are symmetrized with respect to equivalent nuclear pairs (see Sec. II B) and scaled to dimensionless variables,

$$[T^S_{ij}]^{2,0}_{\mathrm{LAB}}(\omega,\Phi) = \left(\frac{2}{3}\right)^{1/2} \frac{4\pi^2 \min\{r_{ij}^3\}}{h\gamma_i\gamma_j} \frac{1}{n_P} \sum_{P=1}^{n_P} \times [T_{\{ij\}_P}]^{2,0}_{\mathrm{LAB}}(\omega,\Phi), \tag{49}$$

where $\{ij\}_P$ is one of the $n_P$ equivalent pairs and $\min\{r_{ij}^3\}$ is the minimum attainable distance for any of these pairs. The standard application of the method has been to proton–proton dipolar couplings but this has a strong limitation because the spectral complication increases with the number of coupled protons. An important development is the possibility of using low ordered lyotropic solvents to limit the number of observable couplings.[5] Near magic angle spinning has also been proposed as a way of reducing dipolar couplings in a controlled way and obtaining order parameters for large molecules in liquid crystals.[50] These methods have opened the application of conformational investigations based on dipolar coupling observables to proteins and other macromolecules. Recently the possibility of studying CH dipolar couplings has also been practically demonstrated.[2] Apart from this practical limitation on the number of couplings, another limit of $D_{ij}$ (and of the $\Delta\nu_k$) is that they vanish in isotropic solution. However, quadrupolar and dipolar couplings are by no means the only NMR observables sensitive to internal motions. In what follows we examine two more.

### C. NMR vicinal spin–spin couplings

The vicinal nuclear spin–spin coupling $^3\mathrm{J}_{ij}$ between a pair of nuclei $i$ and $j$ separated by three chemical bonds, depends on the dihedral angle $\theta$ implicitly defined by these bonds. A semiempirical relation was first introduced by Karplus[51] as

$$^3\mathrm{J}_{ij}(\theta) = a\cos^2\theta + b\cos\theta + c, \tag{50}$$

where $a$, $b$ and $c$ are empirical constants. Similar expressions were developed for the H–C–C–H group, including a correction due to the electronegativity of carbon substituents.[52] The dihedral angle $\theta$ is related to the conformation $\phi_k$ through simple geometrical transformations. In particular, for alkyl chains they simply differ by a phase factor. In general, the averaged observable coupling constant can be written as

$$^3\mathrm{J}_{ij} \equiv \langle [^3\mathrm{J}_{ij}] \rangle = \sum_{n=-2}^{2} [^3\mathrm{J}_{ij}]_n \langle e^{in\phi_k} \rangle = \sum_{n=-2}^{2} [^3\mathrm{J}_{ij}]_n p^{0*}_{0,0;0,\ldots,n,\ldots,0}, \tag{51}$$

where the components $[^3\mathrm{J}_{ij}]_n$ contain all the empirical constants and the phase difference between $\theta$ and $\phi_k$. Notice the direct dependence of $^3\mathrm{J}_{ij}$ on the internal order parameters. Finally, the vicinal spin–spin coupling constants are symmetrized with respect to the exchange of equivalent nuclear pairs, giving

$$[^3\mathrm{J}^S_{ij}](\Phi) = \frac{1}{\max\{|[^3\mathrm{J}_{ij}](\Phi)|\}} \frac{1}{n_P} \sum_{P=1}^{n_P} [^3\mathrm{J}_{\{ij\}_P}](\Phi). \tag{52}$$

This type of measurement can provide information for molecules in isotropic solution and on the purely internal conformation distribution not accessible with the two previous techniques.

### D. Nuclear Overhauser effect

A nuclear Overhauser effect (NOE) is a change in the intensity of a certain observed NMR line for a nucleus $i$ occurring when the population of another nucleus $j$ is satu-

rated or inverted by suitable irradiation.[7,8,10,11,53] Here we consider only a simple subsystem of two 1/2 spins to keep the equations manageable. We assume the coupling between the two nuclei to occur with a dipole–dipole mechanism (cf. Sec. III B) so that the dynamic spin Hamiltonian is

$$\mathcal{H}'_{\rm dip}(t)=\sum_{i<j}\sum_{m=-2}^{2} I_{ij}^{2,m*}\{[T_{ij}]^{2,m}_{\rm LAB}(\omega_t,\Phi_t) - \langle [T_{ij}]^{2,0}_{\rm LAB}(\omega,\Phi)\rangle_{\omega,\Phi}\delta_{m,0}\}, \tag{53}$$

with the sum running on independent nuclear pairs $ij$ and with $I_{ij}^{2,m}$ and $[T_{ij}]^{2,m}_{\rm LAB}$ the spherical components of the nuclear spin operator and the dipole tensor $\mathbf{T}_{ij}$ already defined by Eq. (44). The $\delta_{m,0}$ in the second term comes from the assumed uniaxiality of the liquid crystal solvent around the laboratory $Z$ axis. The standard, motional narrowing, Redfield type relaxation treatment[53,54] leads to the expression for the NOE effect in terms of the spectral densities $J_{ij}^{(m)}(\hat{\omega})$ defined as

$$J_{ij}^{(m)}(\hat{\omega})=4\pi^2\int_0^\infty dt e^{i\hat{\omega}t}C_{ij}^{(m)}(t), \tag{54}$$

i.e., as Fourier transforms at frequency $\hat{\omega}$ of the orientation–conformation dependent dipole–dipole correlation functions,

$$\begin{aligned}
C_{ij}^{(m)}(t)&=\langle [T_{ij}]^{2,m}_{\rm LAB}(\omega_0,\Phi_0)[T^*_{ij}]^{2,m}_{\rm LAB}(\omega_t,\Phi_t)\rangle-\langle [T_{ij}]^{2,0}_{\rm LAB}(\omega,\Phi)\rangle_{\omega,\Phi}\langle [T^*_{ij}]^{2,0}_{\rm LAB}(\omega,\Phi)\rangle_{\omega,\Phi}\delta_{m,0}\\
&=\sum_{n_a,n_b=-2}^{2}\{\langle D^{2*}_{m,n_a}(\omega_0)[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)D^2_{m,n_b}(\omega_t)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_t)\rangle\\
&\quad-\langle D^{2*}_{0,n_a}(\omega)[T_{ij}]^{2,n_a}_{M_1}(\Phi)\rangle_{\omega,\Phi}\langle D^2_{0,n_b}(\omega)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi)\rangle_{\omega,\Phi}\delta_{m,0}\}\\
&=\frac{3}{2}\left(\frac{h\gamma_i\gamma_j}{4\pi^2}\right)^2\sum_{n_a,n_b=-2}^{2}\left\{\left\langle\frac{D^{2*}_{m,n_a}(\omega_0)D^{2*}_{n_a,0}[\omega_{ij}(\Phi_0)]}{r^3_{ij}(\Phi_0)}\,\frac{D^2_{m,n_b}(\omega_t)D^2_{n_b,0}[\omega_{ij}(\Phi_t)]}{r^3_{ij}(\Phi_t)}\right\rangle\right.\\
&\quad\left.-\left\langle\frac{D^{2*}_{m,n_a}(\omega)D^{2*}_{n_a,0}[\omega_{ij}(\Phi)]}{r^3_{ij}(\Phi)}\right\rangle_{\omega,\Phi}\left\langle\frac{D^2_{m,n_b}(\omega)D^2_{n_b,0}[\omega_{ij}(\Phi)]}{r^3_{ij}(\Phi)}\right\rangle_{\omega,\Phi}\delta_{m,0},\right\},
\end{aligned} \tag{55}$$

where $r_{ij}(\Phi_t)$, $\omega_{ij}(\Phi_t)$ are the modulus and the orientation of the $ij$ internuclear vector with respect to the molecular frame $M_1$ in the conformation $\Phi_t$ at time $t$, and the average $\langle\ldots\rangle$ is over the joint distribution $P(\omega_0,\Phi_0,0;\omega_t,\Phi_t,t)$, including molecular tumbling, $\omega$, and conformational, $\Phi$, variables. For instance in an AX spin pair,[55,53]

$$\mathrm{NOE}-1=\frac{\gamma_i}{\gamma_j}\,\frac{6J^{(2)}_{ij}(\hat{\omega}_i+\hat{\omega}_j)-J^{(0)}_{ij}(\hat{\omega}_j-\hat{\omega}_i)}{6J^{(2)}_{ij}(\hat{\omega}_i+\hat{\omega}_j)+3J^{(1)}_{ij}(\hat{\omega}_i)+J^{(0)}_{ij}(\hat{\omega}_j-\hat{\omega}_i)+12j^{(1)}_i(\hat{\omega}_i)}, \tag{56}$$

where $\hat{\omega}_i$, $\hat{\omega}_j$ are the Larmor frequencies of nuclei $i$, $j$ and $j_i^{(m)}(\hat{\omega})$ is the so-called random field providing additional relaxation channels.[53] More generally, the NOE effect will be a result of multispin relaxation and procedures based on a full Redfield relaxation matrix $\mathbf{R}$ rather than on simple analytic formulas that might have to be used. Thus to be as general as possible, we concentrate here on the correlation functions (or spectral densities) from which $\mathbf{R}$ can be built. The correlation functions depend on the overall tumbling of the molecule $\omega_t$ and on the internal dynamics $\Phi_t$. A general treatment of this problem is certainly beyond the scope of this work, but we wish to point out the richness of this type of experiment, by considering the correlation functions in some simple limiting cases, where the motions are so relatively fast or slow that they reduce to time independent, static values $C_{ij}^{(m)}$ (respectively the initial or long time value) on the experiment time scale. The spectral density can therefore be approximated as

$$J_{ij}^{(m)}(\hat{\omega})\approx J_{ij}^{(m)}(0)\approx 4\pi^2 C_{ij}^{(m)}\tau_{ij}^{(m)}, \tag{57}$$

where $\tau_{ij}^{(m)}$ is an effective correlation time for the motion of that pair of nuclei.[56] Notice that at this level we are not concerned with the details of the dynamics, which is in itself a problem of significant complication even for simple biaxial rigid molecules.[57] Indeed we expect that suitable ratios of spectral densities for nuclear pairs sharing a similar dynamics could be considered in applications. We now consider several special cases,

(i) Slow orientational tumbling motion, slow internal motion:

$$\begin{aligned}
C_{ij}^{(m)}\approx&\sum_{n_a,n_b=-2}^{2}\{\langle D^{2*}_{m,n_a}(\omega_0)D^2_{m,n_b}(\omega_0)[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\\
&\times[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}-\langle D^{2*}_{0,n_a}(\omega_0)\\
&\times[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\langle D^2_{0,n_b}(\omega_0)\\
&\times[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\delta_{m,0}\}.
\end{aligned} \tag{58}$$

We can now write this equation in terms of order parameters,

as defined in Eq. (11). To do this we couple the Wigner matrices defining the product $[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)$ and expanding the result in Fourier series,

$$
\begin{aligned}
&[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\\
&\quad=(-)^{n_a}\frac{3}{2}\left(\frac{h\gamma_i\gamma_j}{4\pi^2}\right)^2\sum_{L=0}^{4}C(2,2,L;0,0)\\
&\qquad\times C(2,2,L;-n_a,n_b)\frac{D^L_{n_b-n_a,0}[\omega_{ij}(\Phi_0)]}{r^6_{ij}(\Phi_0)}\\
&\quad=\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}[\Psi_{ij}]^{n_a,n_b}_{a_1,\ldots,a_{N-1}}\\
&\qquad\times\exp\{ia_1\phi_1+\ldots+ia_{N-1}\phi_{N-1}\}. \qquad (59)
\end{aligned}
$$

In a similar way we can couple the product of the two orientational Wigner matrices $D^{2*}_{m,n_a}(\omega_0)D^2_{m,n_b}(\omega_0)$. Combining the results, we can define a new function $[\Upsilon_{ij}]^{L,m,n}_{M_1}(\Phi)$ and expand it in Fourier series,

$$
\begin{aligned}
&\sum_{n_a,n_b=-2}^{2}D^{2*}_{m,n_a}(\omega_0)D^2_{m,n_b}(\omega_0)[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\\
&\quad=\sum_{L=0}^{4}\sum_{n=-L}^{L}D^{L*}_{0,n}(\omega_0)[\Upsilon_{ij}]^{L,m,n}_{M_1}(\Phi_0), \qquad (60)
\end{aligned}
$$

$$
\begin{aligned}
[\Upsilon_{ij}]^{L,m,n}_{M_1}(\Phi)&=\sum_{n_b=-2}^{2}(-)^{m-n_b}C(2,2,L;-m,m)\\
&\quad\times C(2,2,L;-n-n_b,n_b)[\Psi_{ij}]^{n+n_b,n_b}_{M_1}(\Phi)\\
&=\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}[\Upsilon_{ij}]^{L,m,n}_{a_1,\ldots,a_{N-1}}\\
&\quad\times\exp\{ia_1\phi_1+\ldots+ia_{N-1}\phi_{N-1}\}. \qquad (61)
\end{aligned}
$$

Finally, the correlation functions can be written by using the order parameters as

$$
\begin{aligned}
C^{(m)}_{ij}\approx&\sum_{L=0}^{4}\sum_{n=-L}^{L}\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}p^{L*}_{0,n;a_1,\ldots,a_{N-1}}\\
&\times[\Upsilon_{ij}]^{L,m,n}_{a_1,\ldots,a_{N-1}}\\
&-\left|\sum_{n_a=-2}^{2}\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}p^{2*}_{0,n_a;a_1,\ldots,a_{N-1}}\right.\\
&\left.\times[T_{ij}]^{2,n_a}_{a_1,\ldots,a_{N-1}}\right|^2\delta_{m,0}. \qquad (62)
\end{aligned}
$$

In the further special case of negligible rotational–conformational coupling,

$$
\begin{aligned}
C^{(m)}_{ij}\approx&\sum_{n_a,n_b=-2}^{2}\{\langle D^{2*}_{m,n_a}(\omega_0)D^2_{m,n_b}(\omega_0)\rangle_{\omega_0}\\
&\times\langle[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}\\
&-\langle D^{2*}_{0,n_a}(\omega_0)\rangle_{\omega_0}\langle[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\Phi_0}\langle D^2_{0,n_b}(\omega_0)\rangle_{\omega_0}\\
&\times\langle[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}\delta_{m,0}\}\\
=&\sum_{L=0}^{4}\sum_{n=-L}^{L}\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}p^{L*}_{0,n;0,\ldots,0}\\
&\times p^{0*}_{0,0;a_1,\ldots,a_{N-1}}[\Upsilon_{ij}]^{L,m,n}_{a_1,\ldots,a_{N-1}}\\
&-\left|\sum_{n_a=-2}^{2}\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}p^{2*}_{0,n_a;0,\ldots,0}\right.\\
&\left.\times p^{0*}_{0,0;a_1,\ldots,a_{N-1}}[T_{ij}]^{2,n_a}_{a_1,\ldots,a_{N-1}}\right|^2\delta_{m,0}. \qquad (63)
\end{aligned}
$$

This expression, to our knowledge a new result, shows the wealth of orientational–conformational order parameters obtainable from the experiment. In an isotropic phase the equation reduces to

$$
C^{(m)}_{ij}\approx\left(\frac{3}{10}\right)\left(\frac{h\gamma_i\gamma_j}{4\pi^2}\right)^2\left\langle\frac{1}{r^6_{ij}}\right\rangle_{\Phi_0}, \qquad (64)
$$

where we have not indicated explicitly the dependence on conformation $\Phi$ for $r_{ij}$ and $\omega_{ij}$ since no confusion can arise. This expression is equivalent to that obtained by Tropp[56] and reduces to the usual ones for a rigid molecule[7] if the two nuclei belong to the same fragment.

(ii) Slow tumbling motion, fast internal motion:

$$
\begin{aligned}
C^{(m)}_{ij}\approx&\sum_{n_a,n_b=-2}^{2}\{\langle D^{2*}_{m,n_a}(\omega_0)D^2_{m,n_b}(\omega_0)\\
&\times\langle[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\langle[T^*_{ij}]^{2,n_b}_{M_1}(\Phi)\rangle_{\omega_0,\Phi_0}\rangle_{\omega_0}\\
&-\langle D^{2*}_{0,n_a}(\omega_0)[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\langle D^2_{0,n_b}(\omega_0)\\
&\times[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\delta_{m,0}\}. \qquad (65)
\end{aligned}
$$

In the limit of negligible rotational–conformational coupling,

$$C_{ij}^{(m)} \approx \sum_{n_a,n_b=-2}^{2} \{\langle D^{2*}_{m,n_a}(\omega_0) D^{2}_{m,n_b}(\omega_0)\rangle_{\omega_0} - \langle D^{2*}_{0,n_a}(\omega_0)\rangle_{\omega_0}\langle D^{2}_{0,n_b}(\omega_0)\rangle_{\omega_0}\delta_{m,0}\}\langle [T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\Phi_0}\langle [T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}$$

$$= \sum_{n_a,n_b=-2}^{2} \sum_{L=0}^{4} \sum_{\substack{a_1,\ldots,a_{N-1}\\ b_1,\ldots,b_{N-1}=-\infty}}^{\infty} (-)^{m-n_b} C(2,2,L;m,-m) C(2,2,L;n_a,-n_b)$$

$$\times p^{L*}_{0,n_a-n_b;0,\ldots,0} p^{0*}_{0,0;a_1,\ldots,a_{N-1}} p^{0}_{0,0;b_1,\ldots,b_{N-1}} [T_{ij}]^{2,n_a}_{a_1,\ldots,a_{N-1}} [T^*_{ij}]^{2,n_b}_{b_1,\ldots,b_{N-1}}$$

$$- \left| \sum_{n_a=-2}^{2} \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} p^{2*}_{0,n_a;0,\ldots,0} p^{0*}_{0,0;a_1,\ldots,a_{N-1}} [T_{ij}]^{2,n_a}_{a_1,\ldots,a_{N-1}} \right|^2 \delta_{m,0}, \tag{66}$$

where the components $[T_{ij}]^{2,m}_{a_1,\ldots,a_{N-1}}$ have been introduced in Eq. (47). This is the case, for example, of a slowly tumbling macromolecule (i.e., with nearly fixed orientation) with some rapid conformation motions. In an isotropic phase this equation reduces to

$$C_{ij}^{(m)} \approx \frac{3}{10}\left(\frac{h\gamma_i\gamma_j}{4\pi^2}\right)^2 \left[ \left| \left\langle \frac{D^2_{0,0}(\omega_{ij})}{r^3_{ij}} \right\rangle_{\Phi_0} \right|^2 + 2\left| \left\langle \frac{D^{2*}_{1,0}(\omega_{ij})}{r^3_{ij}} \right\rangle_{\Phi_0} \right|^2 + 2\left| \left\langle \frac{D^{2*}_{2,0}(\omega_{ij})}{r^3_{ij}} \right\rangle_{\Phi_0} \right|^2 \right], \tag{67}$$

a case also treated by Tropp.[56]

(iii) Fast tumbling motion, slow internal motion: This could be the case of a small molecule that undergoes some relatively slow conformational change compared to the overall fast tumbling and

$$C_{ij}^{(m)} \approx \sum_{n_a,n_b=-2}^{2} \{\langle\langle D^{2*}_{m,n_a}(\omega_0)\rangle_{\omega_0}\langle D^{2}_{m,n_b}(\omega_0)\rangle_{\omega_0}$$

$$\times [T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}$$

$$- \langle D^{2*}_{0,n_a}(\omega_0)[T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}$$

$$\times \langle D^{2}_{0,n_b}(\omega_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\omega_0,\Phi_0}\delta_{m,0}\}. \tag{68}$$

In the limit of negligible rotational–conformational coupling,

$$C_{ij}^{(m)} \approx \sum_{n_a,n_b=-2}^{2} \{\langle [T_{ij}]^{2,n_a}_{M_1}(\Phi_0)[T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}$$

$$- \langle [T_{ij}]^{2,n_a}_{M_1}(\Phi_0)\rangle_{\Phi_0}\langle [T^*_{ij}]^{2,n_b}_{M_1}(\Phi_0)\rangle_{\Phi_0}\delta_{m,0}\}$$

$$\times \langle D^{2*}_{m,n_a}(\omega_0)\rangle_{\omega_0}\langle D^{2}_{m,n_b}(\omega_0)\rangle_{\omega_0}$$

$$= \sum_{n_a,n_b=-2}^{2} \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} p^{2*}_{m,n_a;0,\ldots,0}$$

$$\times p^{2}_{m,n_b;0,\ldots,0} p^{0*}_{0,0;a_1,\ldots,a_{N-1}} [\Psi_{ij}]^{n_a,n_b}_{a_1,\ldots,a_{N-1}}$$

$$- \left| \sum_{n_a=-2}^{2} \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} p^{2*}_{0,n_a;0,\ldots,0} \right.$$

$$\left. \times p^{0*}_{0,0;a_1,\ldots,a_{N-1}} [T_{ij}]^{2,n_a}_{a_1,\ldots,a_{N-1}} \right|^2 \delta_{m,0}. \tag{69}$$

Notice that this type of contribution vanishes in isotropic liquids but not in a liquid crystal, even if this has not been exploited until now.

In the limiting cases that we have shown, the spectral densities can be viewed in terms of averages over the unknown single particle orientational–conformational distribution function $P(\omega,\Phi)$ of a suitable combination of the Wigner–Fourier basis function [cf. Eq. (8)]. We realize that practical application of NOE in this context might be difficult, especially in separating out the dynamics, but the experiment is particularly rich in information and worth exploring.

### E. Dielectric constant and electric dipole moment

The dependence of permanent electric dipole moment $\boldsymbol{\mu}$ on the intramolecular rotations has been one of the first tools employed to study molecular conformations.[58,12] The fluid phase observables related to $\boldsymbol{\mu}$ and to the square dipole moment tensor $\mathcal{M} \equiv \boldsymbol{\mu}\otimes\boldsymbol{\mu}$ are, in general, the components of second rank dielectric tensor $\boldsymbol{\epsilon}$. In particular, for an anisotropic uniaxial solution of molecules with moment $\boldsymbol{\mu}$, and using the continuum theory of a particle in an ellipsoidal cavity surrounded by an infinite dielectric in presence of a low intensity electric field, it is possible to write[59,60]

$$\frac{\boldsymbol{\epsilon}-\boldsymbol{\epsilon}^{\infty}}{4\pi\rho} = \frac{\boldsymbol{\epsilon}}{\boldsymbol{\epsilon}-\mathbf{n}[\boldsymbol{\epsilon}-\boldsymbol{\epsilon}^{\infty}]} \frac{\langle\mathcal{M}\rangle}{k_B T}, \tag{70}$$

where $\boldsymbol{\epsilon}^{\infty}$ is the dielectric tensor for a field at high frequency, $\rho = N/V$ is the number of molecules per unit volume, $k_B$ is the Boltzmann constant, $T$ is the absolute temperature and $\mathbf{n}$ is the depolarization tensor for the cavity. When the electric field is parallel or perpendicular to the phase director it is possible to compute from dielectric measurements the two components $\langle\mathcal{M}_{\parallel}\rangle$ and $\langle\mathcal{M}_{\perp}\rangle$ of the square dipole moment. The relation between Cartesian and spherical representation is

$$\langle \mathcal{M}^{0,0}_{\mathrm{LAB}}\rangle = -(\tfrac{1}{3})^{1/2}(\langle\mathcal{M}_{\parallel}\rangle + 2\langle\mathcal{M}_{\perp}\rangle), \tag{71}$$

$$\langle \mathcal{M}^{2,0}_{\mathrm{LAB}}\rangle = (\tfrac{2}{3})^{1/2}(\langle\mathcal{M}_{\parallel}\rangle - \langle\mathcal{M}_{\perp}\rangle). \tag{72}$$

The spherical components referred to laboratory frame are related to those measured with respect to system $M_1$ as

$$\mathcal{M}^{2,0}_{\mathrm{LAB}}(\omega,\Phi)=\sum_{n=-2}^{2} D^{2*}_{0,n}(\omega)\mathcal{M}^{2,n}_{M_1}(\Phi). \tag{73}$$

Since the dipole moment $\boldsymbol{\mu}$ and its square $\mathcal{M}$ are global molecular properties they are invariant under molecular symmetry operations. The renormalization is performed dividing each function $\mathcal{M}^{2,0}_{\mathrm{LAB}}$ by the square of the maximum value taken by the dipolar moment over the angular space $[\omega,\Phi]$. This calculation can be simply expressed if the use of the so-called *additivity* hypothesis of bond or fragment dipole moments is justified. The overall molecular dipole moment is then computed as a linear combination of group dipoles, localized in the chemical bonds or in some of the constituent molecular groups, $\boldsymbol{\mu}=\Sigma_i\boldsymbol{\mu}_i$. A large body of data on these group dipole moments has appeared in the literature.[12] With the additivity hypothesis we can finally write the relation between the average square dipole moment measured in the LAB frame and the order parameters

$$\begin{aligned}\langle \mathcal{M}^{2L,0}_{\mathrm{LAB}}\rangle=&\sum_{n=-2L}^{2L}\sum_{q=-1}^{1} C(1,1,2L;q,n-q)\sum_{i,j=1}^{N}\sum_{\substack{a_1^{(i)},\ldots,a_{i-1}^{(i)}\\ b_1^{(i)},\ldots,b_{i-1}^{(i)}=-1}}\sum_{\substack{a_1^{(j)},\ldots,a_{j-1}^{(j)}\\ b_1^{(j)},\ldots,b_{j-1}^{(j)}=-1}}^{1}[\mu_i]^{1,a_1^{(i)}}_{M_i}[\mu_j]^{1,a_1^{(j)}}_{M_j}\\ &\times p^{2L*}_{0,n;b_1^{(i)}+b_1^{(j)},\ldots,b^{(i)}_{\min\{i,j\}}+b^{(j)}_{\min\{i,j\}},b^{(\max\{i,j\})}_{\min\{i,j\}+1},\ldots,b^{(\max\{i,j\})}_{\max\{i,j\}},\ldots,0}G^{1*}_{q,a_2^{(i)},\ldots,a_{i-1}^{(i)},b_1^{(i)},\ldots,b_{i-1}^{(i)},a_1^{(i)}}\\ &\times G^{1*}_{n-q,a_1^{(j)},\ldots,a_{j-1}^{(j)},b_1^{(j)},\ldots,b_{j-1}^{(j)},a_1^{(j)}},L=0,1\end{aligned} \tag{74}$$

where $[\mu_i]^{1,a_1^{(i)}}_{M_i}$ represents the spherical component of the dipole moment tensor of the fragment $i$ measured in the correspondent frame $M_i$, $C(1,1,2L;q,n-q)$ are the Clebsch–Gordan coefficients and the $G$ coefficients are defined in Eq. (13). We can see that the average square dipole moments only depend on a closed set of orientational–conformational order parameters. It is worth noting that the apparently unwieldy Eq. (74) is actually very simple to implement numerically and to code.

### F. Neutron scattering

Scattering techniques are very powerful tools and they provide structural information at both the molecular and structural level. Furthermore, they allow the determination of orientational order parameters of rank higher than four, usually not accessible using other experimental techniques. Here we consider as an example the coherent contribution to the total neutron scattering due to a single molecule (*intramolecular* scattering) and show that it can be analyzed using the maximum entropy approach to provide information on the orientational–conformational distribution. The *intramolecular* contribution is experimentally determined taking advantage of the different scattering lengths of H and D isotopes by comparison of scattering data measured for different mixtures of the fully protonated and the fully deuterated forms of a molecule (see, for example, Ref. 61). The coherent differential cross section $(d\sigma/d\Omega)$ for the elastic neutron scattering of a single molecule containing $N_a$ nuclei is given by the Debye formula referred to laboratory frame

$$\left\langle \frac{d\sigma(\mathbf{Q})}{d\Omega}\right\rangle=\sum_{i,j=1}^{N_a} b_i b_j\langle\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij})\rangle, \tag{75}$$

where $\sigma$ is the scattering probability, $\Omega$ is a solid angle, $b_i$ is the coherent scattering length for the $i$–th nucleus and $\mathbf{r}_{ij}=\mathbf{r}_j-\mathbf{r}_i\equiv r_{ij}\hat{\mathbf{r}}_{ij}$ is the internuclear vector of length $r_{ij}$. The scattering vector $\mathbf{Q}$ has modulus $Q=4\pi\sin\theta/\lambda$, $\lambda$ being the neutron wavelength and $2\theta$ the full scattering angle. The complete static orientational and conformational information is contained in the structure factor $\langle\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij})\rangle$. Using the Rayleigh expansion

$$\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij})=\sum_{L=0}^{\infty} i^L(2L+1)j_L(Qr_{ij})P_L(\hat{\mathbf{Q}}\cdot\hat{\mathbf{r}}_{ij}), \tag{76}$$

where $j_L$ are Bessel functions of fractional order and $P_L$ are Legendre polynomials, it is possible to link the cross section to the molecular orientational–conformational state $(\omega,\Phi)$. In particular, for uniaxial phases only even ranks contribute and we have

$$\begin{aligned}\frac{d\sigma(\mathbf{Q})}{d\Omega}\equiv&\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]_{\mathrm{LAB}}(\omega,\Phi)\\ =&\sum_{i,j=1}^{N_a} b_i b_j\sum_{L=0}^{\infty}\sum_{n=-2L}^{2L}(-)^L(4L+1)\\ &\times P_{2L}(\cos\theta)j_{2L}[Qr_{ij}(\Phi)]\\ &\times D^{2L*}_{0,n}(\omega)D^{2L*}_{n,0}[\omega_{r_{ij}}(\Phi)]\\ =&\sum_{L=0}^{\infty}\sum_{n=-2L}^{2L} D^{2L*}_{0,n}(\omega)\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]^{2L,n}_{M_1}(\Phi),\end{aligned} \tag{77}$$

where $\theta$ is now the angle between scattered beam and phase director, and the Euler angles $\omega_{r_{ij}}$ define the orientation of $\mathbf{r}_{ij}$ with respect to frame $M_1$. The functions $[(d\sigma(Q,\theta)/d\Omega)]^{2L,n}_{M_1}(\Phi)$ are then defined as

$$\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]^{2L,n}_{M_1}(\Phi) = (-)^L(4L+1)P_{2L}(\cos\,\theta) \times \sum_{i,j=1}^{N_a} b_i b_j j_{2L}[Qr_{ij}(\Phi)]D^{2L*}_{n,0}[\omega_{r_{ij}}(\Phi)]. \quad (78)$$

We can write the Fourier expansion of these functions,

$$\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]^{2L,n}_{M_1}(\Phi) = \sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty}\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]^{2L,n}_{a_1,\ldots,a_{N-1}} \times \exp\{ia_1\phi_1+\cdots+ia_{N-1}\phi_{N-1}\}, \quad (79)$$

and thus we can find the relation between the average cross section and an infinite set of order parameters,

$$\left\langle\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]_{\mathrm{LAB}}\right\rangle = \sum_{L=0}^{\infty}\sum_{n=-2L}^{2L}\sum_{a_1,\ldots,a_{N-1}=-\infty}^{\infty} p^{2L*}_{0,n;a_1,\ldots,a_{N-1}} \times\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]^{2L,n}_{a_1,\ldots,a_{N-1}}. \quad (80)$$

Considering the symmetrization scheme of Sec. II B we notice that the cross section defined by Eq. (75) is essentially an average over all the pairs of atoms in the molecule. Since any symmetry operation can be interpreted as an exchange of positions $\mathbf{r}_i$ of equivalent atoms, the cross sections are invariant under application of the projection operators of Sec. II B. The maximum absolute value of the cross section used to normalize this observable is easily obtained, leading to

$$\left[\frac{d\sigma(Q,\theta)^S}{d\Omega}\right]_{\mathrm{LAB}}(\omega,\Phi) = \left(\sum_{i=1}^{N_a} b_i\right)^{-2}\left[\frac{d\sigma(Q,\theta)}{d\Omega}\right]_{\mathrm{LAB}}(\omega,\Phi). \quad (81)$$

## IV. TEST RESULTS

One of the ubiquitous problems of conformational data analysis is to make sure that the method employed can return the conformational information if and only if it is contained in the experimental data set. In other words, a method should ideally have some self–test capability. Here we wish to show how the maximum entropy method can check its capability to actually recover the structural information contained in experimental data, and we present the results of a series of maximum entropy analyses that we have performed on average values of NMR and dipole observable functions generated from a known distribution function. The purpose of these tests was to examine the benefit, if any, of the maximum entropy combined techniques approach. In particular, we have chosen to combine $^1$HNMR (dipolar couplings), $^2$HNMR (quadrupolar splittings) and dielectric, electric dipole (ED) observables and to examine the resulting synergies. In practice, the test entails the simulation of observable data $\langle[\mathcal{F}_I^S]_{\mathrm{LAB}}\rangle$ starting from a known distribution function $P_s(\omega,\Phi)$, followed by the maximum entropy analysis of these pseudoexperimental data to see if a faithful distribution function is retrieved. Notice that the problem is numerically far from trivial but clearly represents a necessary step for the analysis to be meaningful. The input distribution corresponds to a certain set of orientational order parameters $p^L_{m,n;0,\ldots,0}$ for fragment $M_1$ and a given conformational distribution and, for simplicity, we suppose $P_s(\omega,\Phi)$ to be factorized in a purely orientational part, depending on $\omega$, and a purely conformational part, depending on $\Phi$. In practice, we choose

$$P_s(\omega,\Phi) = \frac{1}{Z_0}P_s(\Phi)\exp\left\{\sum_{L,m,n} a^L_{m,n}D^L_{m,n}(\omega)\right\}. \quad (82)$$

For the *discrete* case, we take instead

$$P_s(\omega,\mathbf{j}) = \frac{1}{Z_0}P_s(\mathbf{j})\exp\left\{\sum_{L,m,n} a^L_{m,n}D^L_{m,n}(\omega)\right\}. \quad (83)$$

Since alkyl chains are a prototype structure for conformational problems we have considered here three substituted alkanes that are representative of a class of molecules that could be studied using the combined techniques, namely 1,4–di–bromo–butane (DBB); 1,6–di–bromo–hexane (DBE) and 1,8–di–bromo–octane (DBO). The most severe test for an analysis technique like maximum entropy, which starts from a flat distribution, is to assume the conformational part to represent a single conformer, since it would require an infinite number of observable basis functions to reconstruct such a deltalike distribution [cf. Eq. (7)]. Thus, the three molecules have been simulated in two rigid conformational states, namely a fully elongated chain (i.e., all–*trans*) and the so-called end–*kink*. For the all–*trans* monorotamer the conformational distribution $P_s(\mathbf{j})$ of Eq. (83) is $P_s(t,\ldots,t)=1$ and zero otherwise, while for the end–*kink* conformer, $P_s(\mathbf{j})$ is nonzero only for $\mathbf{j}=(g^-tg^+t,\ldots,t)$. Following Pines, Ref. 22, the dihedral angle $\phi_g$ for *gauche* conformations has been fixed to $\pm 112.5°$, a slightly different value from the classical RIS angle $\pm 2\pi/3$. The bond distances and angles are: $r_{\mathrm{CC}}=1.53$; $r_{\mathrm{CH}}=1.08$, $r_{\mathrm{CBr}}=1.40$ Å, $\widehat{\mathrm{CCC}}=112$ and $\widehat{\mathrm{HCH}}=109°$. For the terminal groups containing Br we have used a tetrahedral geometry. The orientational part of Eq. (83) has been chosen so that the principal orientational order parameters for the $M_1$ frame are $\langle P_2\rangle=0.60$ and $\langle D^2_{0,2}\rangle=0.05$. The order parameters for each fragment are then $S_{zz}=0.60$ and $S_{\mathrm{CD}}=-0.29$. Using both distributions, and after symmetrization with respect to all equivalent pairs, we have calculated $\Delta\nu_i$ and $D_{ij}$ couplings. Due to the symmetry of the all–*trans* and end–*kink* monorotamers, the square dipole moments, $\langle\mathcal{M}_\parallel\rangle$ and $\langle\mathcal{M}_\perp\rangle$, are zero. All simulated observables are listed for the two conformers in Tables I and II. Going from DBB to DBO, we find, respectively, 2, 3, 4 quadrupolar splittings, and 10, 21, 36 dipolar couplings. Considering the two average dipole moments, the number of observables for each bromo–alkane is then $N_{\mathcal{F}}=14$, 26 and 42. We now turn to the maximum entropy analyses of these simulated observables. The *a priori*

TABLE I. Multitechnique maximum entropy combined analyses (case d) of simulated quadrupolar splittings $\Delta\nu_i$ (in KHz), dipolar couplings $D_{ij}$ (in Hz) and square dipole moments $\langle\mathcal{M}_\parallel\rangle$, $\langle\mathcal{M}_\perp\rangle$ (in Debye) for DBB, DBE and DBO in the all–*trans* conformation (monorotamer). Simulated data have been computed using a distribution [cf. Eq. (83)] giving order parameters $S_{zz}=0.60$ and $S_{\mathrm{CD}}=-0.29$ for each fragment. The root mean square error of ME results, the total number of symmetrized observables $N_{\mathcal{F}}$ and independent linear combinations $N_{\mathcal{G}}$ (dimensionless threshold $\tau=0.01$) are reported. The intrinsic distribution $P_i(\mathbf{j})$ [cf. Eq. (35)] has been defined in terms of a *trans–gauche* energy barrier $E_g=2.1$ kJ/mol, $T=300$ K and using the filter function described in the text for discarding all $(\ldots,g^{\pm}g^{\mp},\ldots)$ sterically forbidden sequences.

| | DBB | | DBE | | DBO | |
|---|---|---|---|---|---|---|
| | Simulated | Analyzed | Simulated | Analyzed | Simulated | Analyzed |
| All–*trans* | $N_{\mathcal{F}}=14$ | $N_{\mathcal{G}}=14$ | $N_{\mathcal{F}}=26$ | $N_{\mathcal{G}}=23$ | $N_{\mathcal{F}}=42$ | $N_{\mathcal{G}}=33$ |
| $\Delta\nu_2$ | −81.63 | −81.63 | −81.63 | −81.48 | −81.67 | −81.68 |
| $\Delta\nu_4$ | −81.58 | −81.58 | −81.58 | −81.47 | −81.58 | −81.58 |
| $\Delta\nu_6$ | ⋯ | ⋯ | −81.58 | −81.52 | −81.58 | −81.57 |
| $\Delta\nu_8$ | ⋯ | ⋯ | ⋯ | ⋯ | −81.58 | −81.58 |
| rms error | | 0.00 | | 0.14 | | 0.01 |
| $D_{2,3}$ | 7726.30 | 7726.30 | 7692.52 | 7693.09 | 7692.52 | 7692.89 |
| $D_{2,4}$ | 566.93 | 566.93 | 559.76 | 548.97 | 559.76 | 559.55 |
| $D_{2,5}$ | 125.63 | 125.63 | 108.86 | 91.45 | 108.87 | 108.46 |
| $D_{2,6}$ | −4377.53 | −4377.53 | −4341.66 | −4333.45 | −4341.66 | −4342.17 |
| $D_{2,7}$ | −1146.03 | −1146.03 | −1142.31 | −1140.41 | −1142.31 | −1143.26 |
| $D_{2,8}$ | −337.06 | −337.06 | −337.06 | −336.42 | −337.06 | −336.80 |
| $D_{2,9}$ | −580.73 | −580.73 | −580.73 | −581.86 | −580.73 | −580.87 |
| $D_{2,10}$ | ⋯ | ⋯ | −547.17 | −542.21 | −547.17 | −544.30 |
| $D_{2,11}$ | ⋯ | ⋯ | −380.14 | −381.92 | −380.14 | −383.41 |
| $D_{2,12}$ | ⋯ | ⋯ | −205.28 | −208.19 | −165.34 | −161.91 |
| $D_{2,13}$ | ⋯ | ⋯ | −164.69 | −162.38 | −206.15 | −206.48 |
| $D_{2,14}$ | ⋯ | ⋯ | ⋯ | ⋯ | −162.57 | −161.75 |
| $D_{2,15}$ | ⋯ | ⋯ | ⋯ | ⋯ | −137.76 | −138.06 |
| $D_{2,16}$ | ⋯ | ⋯ | ⋯ | ⋯ | −86.93 | −88.04 |
| $D_{2,17}$ | ⋯ | ⋯ | ⋯ | ⋯ | −77.28 | −77.92 |
| $D_{4,5}$ | 7760.07 | 7760.07 | 7760.07 | 7763.63 | 7760.07 | 7758.94 |
| $D_{4,6}$ | 573.98 | 573.98 | 573.98 | 574.78 | 573.98 | 574.46 |
| $D_{4,7}$ | 142.05 | 142.05 | 142.05 | 155.90 | 142.04 | 142.37 |
| $D_{4,8}$ | ⋯ | ⋯ | −4413.41 | −4403.78 | −4413.38 | −4413.32 |
| $D_{4,9}$ | ⋯ | ⋯ | −1149.76 | −1148.37 | −1149.76 | −1148.92 |
| $D_{4,10}$ | ⋯ | ⋯ | −337.54 | −336.54 | −337.54 | −337.16 |
| $D_{4,11}$ | ⋯ | ⋯ | −582.54 | −580.39 | −582.54 | −582.66 |
| $D_{4,12}$ | ⋯ | ⋯ | ⋯ | ⋯ | −551.67 | −550.07 |
| $D_{4,13}$ | ⋯ | ⋯ | ⋯ | ⋯ | −382.93 | −385.32 |
| $D_{4,14}$ | ⋯ | ⋯ | ⋯ | ⋯ | −166.02 | −165.89 |
| $D_{4,15}$ | ⋯ | ⋯ | ⋯ | ⋯ | −207.03 | −206.27 |
| $D_{6,7}$ | ⋯ | ⋯ | 7760.07 | 7768.17 | 7760.07 | 7760.30 |
| $D_{6,8}$ | ⋯ | ⋯ | 574.10 | 562.11 | 574.10 | 573.48 |
| $D_{6,9}$ | ⋯ | - | 142.40 | 122.00 | 142.40 | 141.59 |
| $D_{6,10}$ | ⋯ | ⋯ | ⋯ | ⋯ | −4413.41 | −4413.13 |
| $D_{6,11}$ | ⋯ | ⋯ | ⋯ | ⋯ | −1149.76 | −1149.22 |
| $D_{6,12}$ | ⋯ | ⋯ | ⋯ | ⋯ | −337.50 | −337.41 |
| $D_{6,13}$ | ⋯ | ⋯ | ⋯ | ⋯ | −582.47 | −581.71 |
| $D_{8,9}$ | ⋯ | ⋯ | ⋯ | ⋯ | 7760.07 | 7759.85 |
| $D_{8,10}$ | ⋯ | ⋯ | ⋯ | ⋯ | 573.98 | 574.38 |
| $D_{8,11}$ | ⋯ | ⋯ | ⋯ | ⋯ | 142.05 | 141.95 |
| rms error | | 0.00 | | 27.07 | | 4.02 |
| $\langle\mathcal{M}_\parallel\rangle^{1/2}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\langle\mathcal{M}_\perp\rangle^{1/2}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| rms error | | 0.01 | | 0.05 | | 0.05 |

distribution function $P_i(\mathbf{j})$ has been defined for the three bromo–alkanes as Eq. (35), using for all rotors a fixed barrier $E_g=2.1$ kJ/mol and $T=300$ K. Sterically forbidden conformations in alkyl chains have been discarded after a suitable filter function $f(\mathbf{j})$ has been implemented.[34] In practice, if the distance $r_{kl}$ between a pair of atoms $\{k,l\}$ belonging to fragments $M_a$, $M_b$ and separated by at least four atoms $(|a-b|\geqslant 4)$ is lower than the sum of their van der Waals radii, $r_{kl}\leqslant r_k^{(v)}+r_l^{(v)}$, the conformation is rejected. We used $r_{\mathrm{C}}^{(v)}=1.7$, $r_{\mathrm{H}}^{(v)}=1.2$, $r_{\mathrm{Br}}^{(v)}=1.7$ Å. No further assumptions, such as the often used united atoms approximation, which collapses a methylene in a single group, is used. In practice, with this filter function all the $(\ldots,g^{\pm}g^{\mp},\ldots)$ sequences are discarded because of steric hindrance, and the number of nonforbidden conformers for DBB, DBE and DBO is $n_T=7$, 99 and 577. The combination and orthogonalization

TABLE II. Multitechnique ME combined analyses (case d) of simulated data for an end–*kink* monorotamer (cf. Table I for additional details).

| | DBB | | DBE | | DBO | |
|---|---|---|---|---|---|---|
| | Simulated | Analyzed | Simulated | Analyzed | Simulated | Analyzed |
| End–*kink* | $N_{\mathcal{F}}=14$ | $N_{\mathcal{G}}=14$ | $N_{\mathcal{F}}=26$ | $N_{\mathcal{G}}=23$ | $N_{\mathcal{F}}=42$ | $N_{\mathcal{G}}=33$ |
| $\Delta\nu_2$ | −81.63 | −81.63 | −81.67 | −81.70 | −81.67 | −81.59 |
| $\Delta\nu_4$ | −2.62 | −2.62 | −42.10 | −42.13 | −42.10 | −42.09 |
| $\Delta\nu_6$ | ⋯ | ⋯ | −42.10 | −42.12 | −42.10 | −42.17 |
| $\Delta\nu_8$ | ⋯ | ⋯ | ⋯ | ⋯ | −81.58 | −81.58 |
| rms error | | 0.00 | | 0.03 | | 0.06 |
| $D_{2,3}$ | 7726.30 | 7726.30 | 7692.52 | 7693.70 | 7692.52 | 7678.87 |
| $D_{2,4}$ | −1142.19 | −1142.19 | −292.86 | −292.77 | −292.85 | −291.81 |
| $D_{2,5}$ | −1435.16 | −1435.16 | −660.87 | −660.46 | −660.85 | −653.14 |
| $D_{2,6}$ | 677.99 | 677.99 | −1833.43 | −1834.09 | −1833.43 | −1835.30 |
| $D_{2,7}$ | 159.19 | 159.19 | −496.26 | −496.19 | −496.26 | −494.82 |
| $D_{2,8}$ | −1.73 | −1.73 | −279.37 | −279.14 | −169.40 | −169.65 |
| $D_{2,9}$ | −22.33 | −22.33 | −243.30 | −243.11 | −301.54 | −304.03 |
| $D_{2,10}$ | ⋯ | ⋯ | −388.77 | −397.38 | −543.02 | −538.09 |
| $D_{2,11}$ | ⋯ | ⋯ | −374.33 | −364.80 | −403.32 | −411.12 |
| $D_{2,12}$ | ⋯ | ⋯ | −122.34 | −125.40 | −145.12 | −138.23 |
| $D_{2,13}$ | ⋯ | ⋯ | −101.05 | −101.83 | −147.46 | −145.98 |
| $D_{2,14}$ | ⋯ | ⋯ | ⋯ | ⋯ | −174.68 | −175.78 |
| $D_{2,15}$ | ⋯ | ⋯ | ⋯ | ⋯ | −163.15 | −163.80 |
| $D_{2,16}$ | ⋯ | ⋯ | ⋯ | ⋯ | −76.07 | −78.58 |
| $D_{2,17}$ | ⋯ | ⋯ | ⋯ | ⋯ | −67.86 | −68.86 |
| $D_{4,5}$ | 704.85 | 704.85 | 4232.46 | 4234.19 | 4232.46 | 4227.51 |
| $D_{4,6}$ | 854.07 | 854.07 | 714.03 | 715.01 | 714.03 | 709.56 |
| $D_{4,7}$ | 2280.49 | 2280.49 | 1211.29 | 1212.30 | 1211.28 | 1206.82 |
| $D_{4,8}$ | ⋯ | ⋯ | −247.82 | −247.50 | −1866.08 | −1870.18 |
| $D_{4,9}$ | ⋯ | ⋯ | −1325.45 | −1325.51 | −490.59 | −490.93 |
| $D_{4,10}$ | ⋯ | ⋯ | −827.60 | −827.91 | −582.57 | −587.12 |
| $D_{4,11}$ | ⋯ | ⋯ | −2775.05 | −2776.17 | −1678.79 | −1675.21 |
| $D_{4,12}$ | ⋯ | ⋯ | ⋯ | ⋯ | −291.15 | −307.69 |
| $D_{4,13}$ | ⋯ | ⋯ | ⋯ | ⋯ | −372.14 | −360.32 |
| $D_{4,14}$ | ⋯ | ⋯ | ⋯ | ⋯ | −373.09 | −398.07 |
| $D_{4,15}$ | ⋯ | ⋯ | ⋯ | ⋯ | −465.08 | −434.37 |
| $D_{6,7}$ | ⋯ | ⋯ | 4232.46 | 4233.95 | 4232.46 | 4234.90 |
| $D_{6,8}$ | ⋯ | ⋯ | −1138.92 | −1137.81 | −282.42 | −279.28 |
| $D_{6,9}$ | ⋯ | ⋯ | −1439.78 | −1439.28 | −648.71 | −645.90 |
| $D_{6,10}$ | ⋯ | ⋯ | ⋯ | ⋯ | −2795.11 | −2792.78 |
| $D_{6,11}$ | ⋯ | ⋯ | ⋯ | ⋯ | −1984.62 | −1985.77 |
| $D_{6,12}$ | ⋯ | ⋯ | ⋯ | ⋯ | −559.78 | −559.68 |
| $D_{6,13}$ | ⋯ | ⋯ | ⋯ | ⋯ | −466.69 | −467.09 |
| $D_{8,9}$ | ⋯ | ⋯ | ⋯ | ⋯ | 7760.07 | 7759.49 |
| $D_{8,10}$ | ⋯ | ⋯ | ⋯ | ⋯ | 574.00 | 574.73 |
| $D_{8,11}$ | ⋯ | ⋯ | ⋯ | ⋯ | 142.09 | 136.39 |
| rms error | | 0.01 | | 9.67 | | 29.22 |
| $\langle\mathcal{M}_{\parallel}\rangle^{1/2}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\langle\mathcal{M}_{\perp}\rangle^{1/2}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| rms error | | 0.01 | | 0.05 | | 0.05 |

algorithm of Sec. II B, with the *intrinsic* distribution $P_i(\mathbf{j})$ as weight factor and a dimensionless threshold $\tau=0.01$, provides $N_{\mathcal{G}}=14$, 23 and 33 independent basis functions for the three bromo–alkanes. We see that, except for DBB, at least two redundant parameters have been eliminated by this orthogonalization procedure. For each bromo–alkane, we have performed four different maximum entropy analyses. The first three considered only one set of observables, namely (a) $^2$HNMR quadrupolar splittings, (b) $^1$HNMR dipolar couplings, and (c) dipolar data. The fourth one (d) was a maximum entropy combined analysis of all three data sets. In all cases, the agreement between input (i.e., pseudoexperimental) observables and recalculated ones was very good [root mean square (rms) error generally less than $\pm 1\%$], demonstrating that the analysis is feasible and sufficiently well conditioned from a numerical point of view also for complex molecules. The recalculated values of $S_{zz}$ and $S_{\mathrm{CD}}$ from the (a), (b) and (d) maximum entropy analyses were excellent

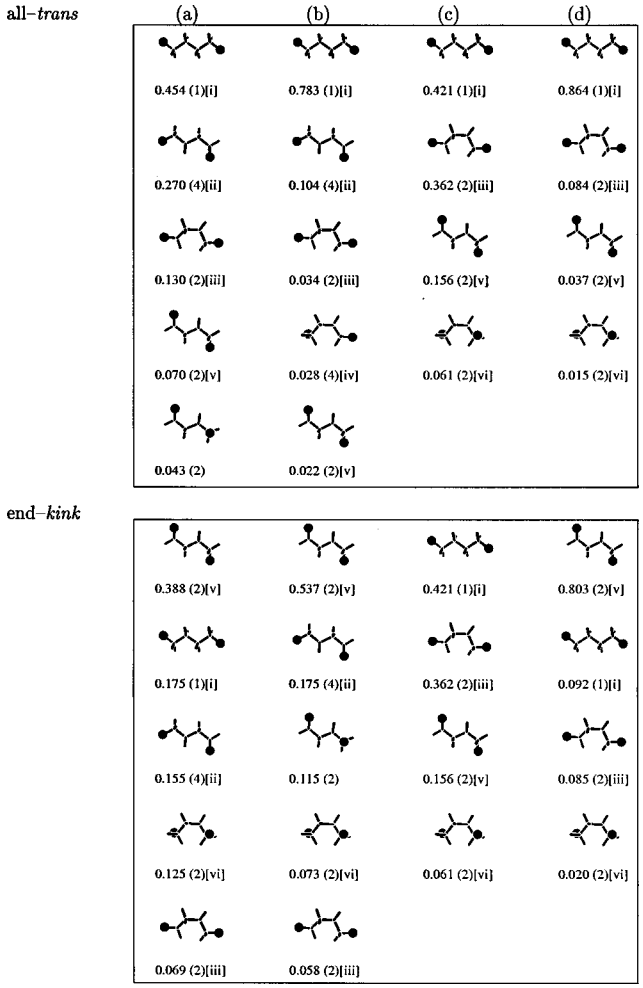


FIG. 3. The four maximum entropy analyses (cases a–d) of simulated all–*trans* (top plate) and end–*kink* (bottom plate) observables for DBB as described in the text. For every analysis we show the six most probable conformers with their total probability $mP(\mathbf{j})$ (the multiplicity $m$ is given between curly brackets). The most frequently occurring conformational sequences are labeled as $[i]\equiv(t,\ldots,t)$, $[ii]\equiv(g^+tt,\ldots,t)$, $[iii]\equiv(tg^+t,\ldots,t)$, $[iv]\equiv(g^+g^+t,\ldots,t)$, $[v]\equiv(g^-tg^+,\ldots,t)$ and $[vi]\equiv(g^+g^+g^+,\ldots,t)$.

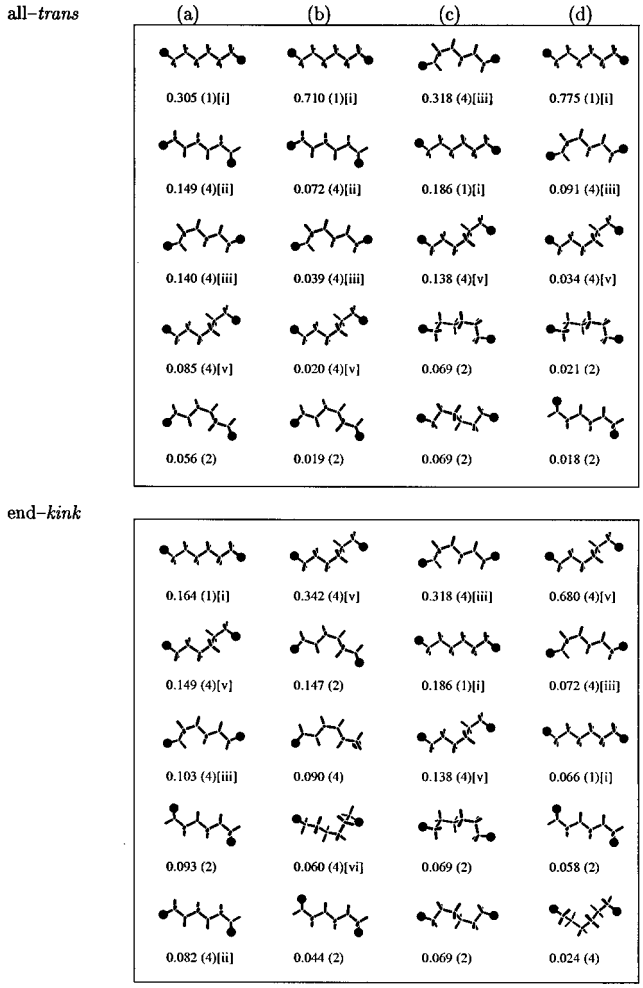


FIG. 4. The four maximum entropy analyses (cases a–d) of simulated all–*trans* (top plate) and end–*kink* (bottom plate) observables for DBE as described in the text (cf. Fig. 3 for additional details).

(error less than 0.1%), especially for the combined case (d), and only the analysis of electric dipole moments (case c) could not recover the input order parameters. In Tables I and II we report the results of analyses (d) for the all–*trans* and end–*kink* simulated data. We now discuss the conformational information recovered from the analysis. Analyzing and summarizing structural information for various molecules is not a trivial task. We have then chosen to present the principal features of the maximum entropy distribution showing the five most probable conformations recovered (see Figs. 3–5) and labeling the relevant sequences. The discrete $\mathbf{j}$ states can be divided in three classes according to their degeneracy $m$, that is: (i) $m=1$ for the all–*trans* conformation, i.e., $\mathbf{j}=(t,\ldots,t)$; (ii) $m=2$ for center–symmetric sequences where each *gauche* state appears in *mirror* sites $j_k$ and $j_{N-k}$, e.g., $\mathbf{j}=(g^+tg^+,\ldots,g^+tg^+)\equiv(g^-tg^-,\ldots,g^-tg^-)$; (iii) $m=4$ for nonsymmetric sequences, e.g., $\mathbf{j}=(t,\ldots,tg^+g^+)\equiv(t,\ldots,tg^-g^-)\equiv(g^+g^+t,\ldots,t)\equiv(g^-g^-t,\ldots,t)$. We use the product $mP(\mathbf{j})$ to measure the relative abundance of each conformational sequence. To simplify the notation, the multiplicity factor has been automatically included in $P(\mathbf{j})$ in the following discussion. In Fig. 3 we report the maximum entropy results for simulated all–*trans* (top) and end–*kink* (bottom) DBB. We see that the four analyses (a–d) of all–*trans* data have all retrieved the *trans* conformer as the most probable. The sharpest distribution corresponds to the combined analysis (d) with $P(t,\ldots,t)=0.864$, closely followed by the $^1$HNMR case (b) with $P(t,\ldots,t)=0.783$. Columns (c) and (d) list only four conformers because the inclusion of practically zero ED data forced the rejection of all sequences with a not zero total dipole moment. Turning now to the maximum entropy analyses of the simulated end–*kink* conformer (Fig. 3, bottom), we see that the most probable conformation recovered for DBB is actually end–*kink* only for cases (a), (b) and (d) and that ED data alone are not sufficient to counterbalance the *intrinsic* bias for elongated configurations accounted for by $P_i(\mathbf{j})$. For a similar reason, the difference between the combined analysis, $P(g^-tg^+t,\ldots,t)=0.803$, and the $^1$HNMR case $P(g^-tg^+t,\ldots,t)=0.537$ is higher than that for analyzed all–*trans* data. Similar conclusions can be drawn from the analyses of simulated data for

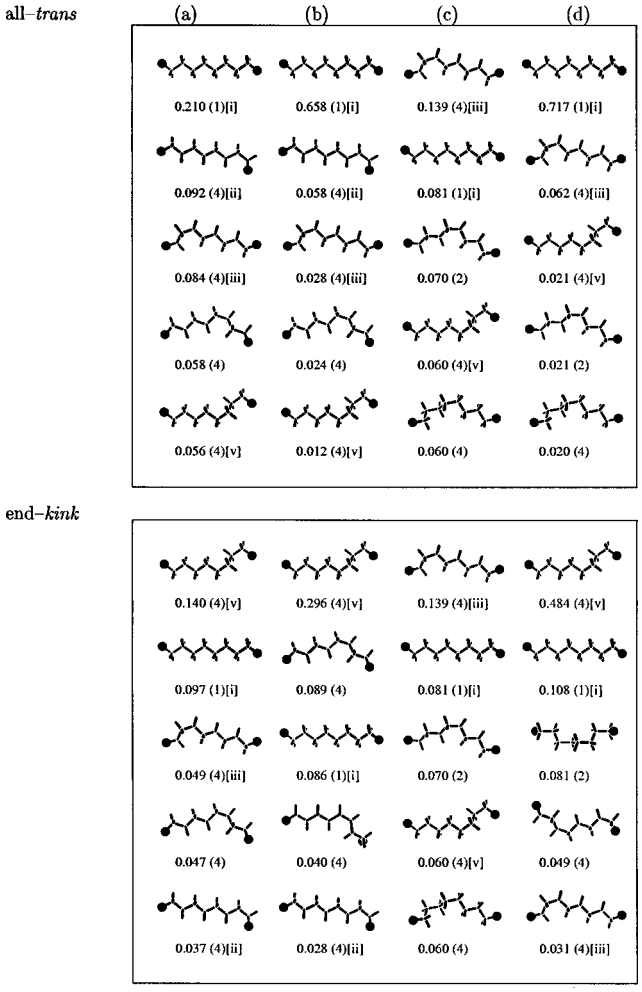


FIG. 5. The four maximum entropy analyses (cases a–d) of simulated all–*trans* (top plate) and end–*kink* (bottom plate) observables for DBO as described in the text (cf. Fig. 3 for additional details).

DBE and DBO (Figs. 4 and 5). Except for the (c) case, the *trans* conformer is the most probable for analyses of simulated all-*trans* DBE (Fig. 4, top). Again, the sharpest result is that of combined case (d), closely followed by the $^1$HNMR (b) case. The synergy of combined techniques is more evident for the end–*kink* observables (Fig. 4, bottom). The retrieved probability almost doubles, $P(g^-tg^+t,\ldots,t) = 0.680$, including $^2$HNMR and ED data to the analysis of $^1$HNMR data which only gives $P(g^-tg^+t,\ldots,t) = 0.342$. For DBO (Fig. 5), the most likely sequences recovered by the combined analyses of simulated *trans* and *kink* data are $P(t,\ldots,t) = 0.717$ and $P(g^-tg^+t,\ldots,t) = 0.484$. All the conformational probability obtained, and in particular the order of the single conformer in the probability scale, do not change by choosing the $E_g$ value in the typical range [2.1–2.5] kJ/mol calculated from experimental studies of butane dissolved in liquid solvents.[62] In any case, the inclusion of the *prior* information to the maximum entropy method by setting the $E_g$ value drastically changes the conformer probability distribution because the intrinsic function $P_i$ determines to a large extent what is obtained. These results are useful to confirm that the maximum entropy method can simultaneously analyze experimental data from different techniques without the use of approximate mean field models to describe flexible molecules in liquid crystal solution.

## V. CONCLUSIONS

It is becoming increasingly clear that no single experimental technique can provide sufficiently detailed data to give a complete representation of the orientational–conformational distribution for a flexible solute in liquid crystals and that reliable conformation information for complex molecules in anisotropic solution, particularly molecules of biological interest, can only come by carefully planning various experiments using different techniques on the same system. Here we have proposed a unified theoretical framework to simultaneously analyze data from different techniques using the maximum entropy method in terms of the least biased overall orientational–conformational distribution. Given a molecule of interest and a basic molecular skeleton, the present technique can, first of all, assess the feasibility of the project by constructing simulated data for certain conformations, adding noise to a desired level if appropriate, and analyzing these pseudoexperimental data as the experimental ones with the help of any prior existing information. Various candidate techniques can be tried until a promising combination of experiments is found. We have demonstrated the method on terminally dibrominated alkanes data considering simulated $^1$HNMR, $^2$HNMR and dielectric observables. Summarizing, we notice that for these molecules, $^1$HNMR data are the most useful, followed by $^2$HNMR and dielectric. Dielectric data alone are not really informative enough to provide a reliable conformational distribution, in particular for the longest alkyl chain (DBO). Nevertheless, combining dielectric data with $^1$HNMR couplings considerably improves the maximum entropy results. This synergistic effect is smaller if we merge $^2$HNMR splittings to $^1$HNMR couplings. Other techniques can be brought into play and we expect NOE data, already very useful in isotropic solution, to be particularly important. We hope the availability of a relatively simple data analysis framework such as the present one will stimulate the combined experiments needed to provide the much needed leap in the investigation of conformational and rotational–conformational distribution in solution.

## ACKNOWLEDGMENTS

We are grateful to the University of Bologna, MURST ex 40%, CNR, EU TMR FMRX CT970121 for support of this work. We also thank Professor E. E. Burnell for comments on the manuscript.

[1] For an overview see, *Faraday Symposium 27 on The Conformation of Flexible Molecules in Fluid Phases*, J. Chem. Soc., Faraday Trans. **88** (1992).
[2] K. Schmidt-Rohr, D. Nanz, L. Emsley, and A. Pines, J. Phys. Chem. **98**, 6668 (1994).
[3] D. Sanström, K. T. Summanenand, and M. H. Levitt, J. Am. Chem. Soc. **116**, 9357 (1994).
[4] J. W. Emsley and J. C. Lindon, *NMR Spectroscopy Using Liquid Crystals Solvents* (Pergamon, Oxford, 1975).
[5] N. Tjandra and A. Bax, Science **278**, 1111 (1997).

[6] D. J. Photinos, E. T. Samulski, and H. Toriumi, J. Phys. Chem. **94**, 4688 (1990); **94**, 4694 (1990).
[7] D. Neuhaus and M. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis* (VCH, New York, 1989).
[8] K. Wüthrich, Acc. Chem. Res. **22**, 36 (1989).
[9] R. Kaptein, R. Boelens, R. M. Scheak, and W. F. von Gunsteren, in *Perspectives in Biochemistry*, edited by H. Neurath (American Chemical Society, Washington, DC, 1989), Vol. I, p. 14.
[10] L. Poppe, J. Am. Chem. Soc. **18**, 8421 (1993).
[11] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon, Oxford, 1987).
[12] R. J. W. Le Fevre, *Dipole Moments*, 3rd ed. (Methuen, London, 1953).
[13] C. J. F. Böttcher and P. Bordewijk, *Theory of Electric Polarization* (Elsevier, Amsterdam, 1978), Vol. I.
[14] S. Marčelja, J. Chem. Phys. **60**, 3599 (1974).
[15] A. Ferrarini, G. J. Moro, P. L. Nordio, and G. R. Luckhurst, Mol. Phys. **77**, 1 (1992).
[16] J. W. Emsley and G. R. Luckhurst, Mol. Phys. **41**, 19 (1980); J. W. Emsley, G. R. Luckhurst, and C. P. Stockley, Proc. R. Soc. London, Ser. A **381**, 117 (1982).
[17] A. J. Van der Est, M. Y. Kok, and E. E. Burnell, Mol. Phys. **60**, 397 (1987).
[18] D. Catalano, C. Forte, C. A. Veracini, and C. Zannoni, Isr. J. Chem. **23**, 283 (1983).
[19] J. P. Straley, Phys. Rev. A **10**, 1881 (1974).
[20] D. S. Zimmerman and E. E. Burnell, Mol. Phys. **69**, 1059 (1990); **78**, 687 (1993).
[21] J. M. Polson and E. E. Burnell, J. Chem. Phys. **103**, 6891 (1995).
[22] (a) M. E. Rosen, S. P. Rucker, C. Schmidt, and A. Pines, J. Phys. Chem. **97**, 3858 (1993); (b) M. E. Rosen, Ph.D. thesis, University of California, Berkeley, 1992.
[23] M. Luzar, M. E. Rosen, and S. Caldarelli, J. Phys. Chem. **100**, 5098 (1996).
[24] R. Tarroni and C. Zannoni, J. Phys. Chem. **100**, 17,157 (1996).
[25] C. Zannoni, in *Nuclear Magnetic Resonance of Liquid Crystals*, edited by J. W. Emsley (Reidel, Dordrecht, 1985), Chap. 2.
[26] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[27] B. J. Berne and G. D. Harp, Adv. Chem. Phys. **17**, 171 (1970).
[28] T. S. Clarkson and G. Williams, J. Chem. Soc., Faraday Trans. 2 **70**, 1705 (1974).
[29] D. I. Bower, J. Polym. Sci., Polym. Phys. Ed. **19**, 93 (1981); Polymer **23**, 1251 (1982).
[30] C. Zannoni, in *Polarized Spectroscopy of Ordered Systems*, edited by B. Samorì and E. Thulstrup (Kluwer, Dordrecht, 1988), p. 57.
[31] L. DiBari, C. Forte, C. A. Veracini, and C. Zannoni, Chem. Phys. Lett. **143**, 263 (1988).
[32] D. Catalano, L. DiBari, C. A. Veracini, N. Shilstone, and C. Zannoni, J. Chem. Phys. **94**, 3928 (1991).
[33] R. Berardi, F. Spinozzi, and C. Zannoni, J. Chem. Soc., Faraday Trans. **88**, 1863 (1992); Liq. Cryst. **16**, 381 (1994).
[34] R. Berardi, F. Spinozzi, and C. Zannoni, Chem. Phys. Lett. **260**, 633 (1996).
[35] P. Flory, *Statistical Mechanics of Chain Molecules* (Wiley Interscience, New York, 1969).
[36] A. J. Dianoux and F. Volino, J. Phys. (Paris) **40**, 181 (1979).
[37] Y. Sasanuma and A. Abe, Polym. J. (Tokyo) **23**, 117 (1991).
[38] M. E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).
[39] W. Maier and A. Saupe, Z. Naturforsch. A **13A**, 564 (1958); **14A**, 882 (1959); **15A**, 287 (1960).
[40] J. Maruani, A. Hernandez Laguna, and Y. G. Smeyers, J. Chem. Phys. **63**, 4515 (1975).
[41] S. L. Altmann, *Induced Representations in Crystals and Molecules* (Academic, London, 1977).
[42] *The Maximum Entropy Formalism*, edited by R. D. Levine and M. Tribus (MIT, Cambridge, MA, 1979).
[43] L. R. Mead and N. Papanicolau, J. Math. Phys. **25**, 2404 (1984).
[44] D. Catalano, J. W. Emsley, G. La Penna, and C. A. Veracini, J. Chem. Phys. **105**, 10595 (1996).
[45] B. R. Frieden, in *Maximum Entropy and Bayesian Methods in Inverse Problems*, edited by C. R. Smiths and W. T. Grandy, Jr. (Reidel, Dordrecht, 1985).
[46] L. R. Pratt and D. Chandler, J. Chem. Phys. **68**, 4202 (1978); **68**, 4213 (1978).
[47] J. Seelig, Q. Rev. Biophys. **10**, 353 (1977).
[48] R. Y. Dong, *Nuclear Magnetic Resonance of Liquid Crystals* (Springer, New York, 1994).
[49] J-P. Douliez, A. Léonard, and E. J. Dufourc, Biophys. J. **68**, 1727 (1995).
[50] A. Kimura, N. Kuni, and H. Fujiwara, J. Phys. Chem. **100**, 14056 (1996).
[51] M. Karplus, J. Chem. Phys. **30**, 11 (1959).
[52] C. A. G. Haasnoot, F. A. A. M. de Leeuw, and C. Altona, Tetrahedron **36**, 2783 (1980).
[53] D. M. Grant, C. L. Mayne, F. Liu, and T. Xiang, Chem. Rev. **91**, 1591 (1991).
[54] A. G. Redfield, Adv. Magn. Reson. **1**, 1 (1965).
[55] J. M. Courtieu, C. M. Mayne, and D. M. Grant, J. Chem. Phys. **66**, 2669 (1977).
[56] J. Tropp, J. Chem. Phys. **72**, 6035 (1980).
[57] R. Tarroni and C. Zannoni, J. Chem. Phys. **95**, 4550 (1991).
[58] C. P. Smyth, in *Internal Rotation in Molecules* (Wiley, London, 1974).
[59] G. R. Luckhurst and C. Zannoni, Proc. R. Soc. London, Ser. A **343**, 389 (1975).
[60] G. Williams, in *The Molecular Dynamics of Liquid Crystals* (Kluwer, Dordrecht, 1994), p. 431.
[61] R. W. Date, I. W. Hamley, G. R. Luckhurst, J. M. Seddon, and R. M. Richardson, Mol. Phys. **76**, 951 (1992).
[62] D. A. Cates and A. MacPhail, J. Phys. Chem. **95**, 2209 (1991).