

Two- and N -step correlated models for the analysis of molecular dynamics trajectories of linear molecules in silicalite

Pierfranco Demontis,^{a)} Giuseppe B. Suffritti, and Antonio Tilocca
Dipartimento di Chimica, Università degli Studi di Sassari, Via Vienna 2, I-07100 Sassari, Italy

(Received 25 April 2000; accepted 2 August 2000)

Recent molecular dynamics data on the diffusion of linear diatomic and triatomic molecules in the zeolite silicalite are analyzed in terms of a new correlated model [F. Jousse, S. M. Auerbach, and D. P. Vercauteren, *J. Chem. Phys.* **112**, 1531 (2000)] capable to account for both first- and higher-order correlation effects. This “ N -step” model reproduces very well our calculated mean square displacements and diffusion coefficients of the molecules considered. The improvements with respect to the results obtained with our previous “two-step” model [P. Demontis, J. Kärger, G. B. Suffritti, and A. Tilocca, *Phys. Chem. Chem. Phys.* **2**, 1455 (2000)] are remarkable for all molecules except chlorine, showing that only in this case the effect of (negative) correlations *spanning more than two jumps* between channel intersections (~ 20 Å) can be neglected. The basic trajectory analysis in terms of single- and two-step models, besides being an useful reference, provides *all* the input data needed for the application of the N -step model. Indeed, in its silicalite formulation, the N -step model is strongly linked to the two-step one because it calculates the probability of a sequence of jumps in the same channel by means of the correlations between any two consecutive jumps. Finally, the possibility to obtain qualitative insight into the diffusive mechanism through various kind of correlation coefficients is discussed. © 2000 American Institute of Physics. [S0021-9606(00)70741-0]

I. INTRODUCTION

The diffusion of guest molecules in zeolites is often conveniently considered as proceeding by uncorrelated hops from site to site.^{1–3} This “random-walk” (RW) picture is useful both to extend the time scale of the simulations,^{4–7} compared to that accessible by standard molecular dynamics (MD) methods,^{8,9} and to gain insight into the diffusion mechanism.^{10–14} In a basic random-walk representation of molecular propagation in zeolite silicalite the molecular motion is resolved into a series of uncorrelated jumps between adjacent channel intersections.¹ When consecutive jumps are not completely uncorrelated, the random-walk picture is no more adequate, and the effects of correlation should be taken into account. We have previously devised a “two-step” model for the molecular diffusion in silicalite, which considerably improves the random-walk description of the diffusion of ethane at different loading and temperature¹⁵ and of linear diatomic and triatomic flexible molecules at infinite dilution.¹⁶ This observation shows that, while the RW uncorrelated picture may be appropriate in the case of small spherical sorbates such as methane and xenon,^{1,17–20} the motion of small linear species in silicalite can exhibit significant correlation effects. The two-step model merges each pair of two successive displacements between channel intersections and represents the molecular motion as a sequence of such double steps. Each two-step jump is assumed to be uncorrelated from the others. This assumption is much more accurate than the random-walk one, but it can still be inadequate for sorbates with long-range diffusive memory where higher-

order correlations could be present. Jousse *et al.*²¹ have recently developed a general model for the diffusion in zeolites; when applied to silicalite this model deals with correlation effects in a more complete and efficient way allowing us to incorporate the residual correlation not included in the two-step model. This is achieved by cutting the molecular motion into an uncorrelated sequences of jumps. Each sequence is made of N consecutive jumps between adjacent intersections in the same channel: accordingly, the model will be termed “ N -step” in the following. The input data needed to compute the N -step diffusion coefficients are easily derived from the probabilities and durations of single- and two-step displacements. Therefore, a suitable way to study the nature and extent of correlation effects in silicalite should in any case proceed by analyzing the MD trajectory via the RW and two-step models. This analysis provides both the data required to compute the diffusion coefficients according to the two corresponding models and those needed to apply the N -step model and to evaluate the diffusion constants according to it. In this article we apply this method to the diffusion of linear diatomic and triatomic molecules representing halogens, carbon dioxide, and carbon disulfide in silicalite at infinite dilution, whose two-step data have recently been reported.¹⁶ As for ethane, the motion of these species shows significant correlation effects. It is important to understand how these effects affect the diffusive motion. The spatial and temporal extent of the diffusive memory, as well as its detailed features (e.g., the presence of positive or negative correlation) are very important parameters both for a qualitative understanding of the diffusive behavior and for improving stochastic diffusion simulation methods which

^{a)}Electronic mail: demontis@ssmain.uniss.it

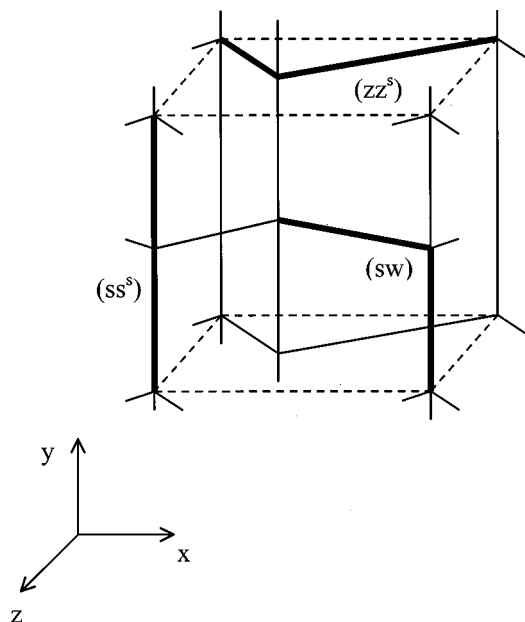


FIG. 1. Schematic representation of the channels geometry in a silicalite unit cell; channels are represented by continuous lines. The three main two-step jumps are shown as thick lines: (ss^s), two displacements in straight channels, in the same y direction; (sw), a switch from a straight to a zigzag channel, or vice versa; (zz^s), two displacements in zigzag channels, in the same x direction.

should include such effects.¹⁴ From this perspective we illustrate the utility of a comparison between diffusion models including correlations up to different orders. The two-step model, incorporating only correlation between two consecutive jumps, is a sort of first-order improvement to the RW picture. However, in the silicalite case, the two-step parameters (i.e., data obtained from a direct analysis of the MD trajectory in terms of the two-step model)¹⁵ are actually the only input data required to the N -step extension²¹ (see Sec. II). In view of that, they appear to hold more information concerning the correlation effects than just first-order effects, so that a two-step analysis is an important starting point for a complete study of correlation effects in the diffusion of small linear molecules in silicalite.

II. CALCULATIONS

The microporous structure of silicalite is made of straight channels intersecting orthogonally with sinusoidal channels both with diameter of about 5.5 Å. The arrangement of channels in a unit cell is schematically depicted in Fig. 1, where the three main two-step jumps are also illustrated. The details of the MD simulations and the analysis of the trajectories in terms of single-step and two-step jumps are reported in Ref. 16, and we will shortly summarize them. The full flexibility of both the zeolite lattice and the guest species was accounted for through effective potentials¹⁶ while the intermolecular interaction between the host framework and the guest molecules was modeled through a Lennard–Jones potential; as only one sorbed molecule is considered in each simulation, no guest–guest interactions are present. For each molecule considered, an MD run of 20 ns was carried out. The MD trajectory provides the “stan-

dard” total diffusion coefficients and its components through the long-time slope of the corresponding mean square displacement (MSD) curves.⁸ Moreover, after mapping the molecular path in terms of the channel intersections visited in succession, we determined the number and duration of “single” steps (i.e., simple jumps between adjacent intersections) and “two-steps” (i.e., displacements composed of two consecutive single steps, crossing *three* intersections). In the RW model the MSD components are associated to the number of single steps in the straight and sinusoidal channels, while the MSD components according to the two-step model can be expressed by means of the calculated numbers of the three types of two-step displacements shown in Fig. 1. The corresponding expressions and the underlying theory are thoroughly described in Ref. 15.

In the N -step model²¹ the focus is on the sequences of N successive jumps (i.e., N consecutive single steps) in the *same channel*: in silicalite the three-dimensional arrangement of orthogonal channels entails that all correlation is lost as soon as the molecule turns from a straight into a sinusoidal channel, or vice versa. Accordingly, if each sequence ends when the molecule switches channel, two successive sequences will be *uncorrelated* from each other. Therefore, by extracting from the molecular trajectory the sequences of N jumps in each channel the motion can be precisely described as a general random walk where each random displacement actually is a N -step sequence made of several single steps within the same channel. The general expression for the MSD along the x axis is then

$$\langle x^2(t) \rangle = N_{\text{seq}}^{\text{zc}}(t) \langle x^2 \rangle_{\text{seq}}, \quad (1)$$

where $N_{\text{seq}}^{\text{zc}}(t)$ is the number of uncorrelated sequences in the zigzag channel observed in the time t , and $\langle x^2 \rangle_{\text{seq}}$ the MSD averaged over such sequences; a similar expression holds for the straight channel. When the results arising from this kind of expressions are more accurate, the less correlation persists between the events they are made from. While in the standard random-walk model in silicalite¹ each event is a single displacement between two adjacent intersections (i.e., $\langle x^2 \rangle_{\text{seq}}$ in Eq. (1) is the square distance between adjacent intersections along the sinusoidal channel), the two-step expressions have been built considering all possible sequences of two single steps,¹⁵ and the corresponding distance covered in each such sequence. As noted above the extension to the N -step model in silicalite is rather direct: it only requires the numbers of two-step events observed in the MD trajectory as well as the duration of single steps in both channels, i.e., data already available from our previous random-walk and two-step analysis. Indeed, as shown by Jousse *et al.*,²¹ by expressing the total number of sequences N_{seq} as a function of the two-step data and noting that, due to the strict alternation of sequences in straight and zigzag channels, $N_{\text{seq}}^{\text{sc}} = N_{\text{seq}}^{\text{zc}} = (1/2)N_{\text{seq}}$ the number of sequences in either channel is given by

$$N_{\text{seq}}^{\text{sc}}(t) = N_{\text{seq}}^{\text{zc}}(t) = \left[\frac{t_s}{p_{sz}} + \frac{t_z}{p_{zs}} \right]^{-1} t, \quad (2)$$

where t_s and t_z are the average time lengths of single steps in the straight and zigzag channel, respectively, while p_{sz} and

p_{zs} are the two-step probabilities of moving in a zigzag channel after a step in a straight one, or vice versa.²² The average MSD during one sequence, in terms of jumps, is:

$$\langle n^2 \rangle_{\text{seq}} = \frac{1+p\chi}{1-p\chi} \left(\frac{1}{1-p} \right), \quad (3)$$

where p is the ‘‘channel probability’’ that the molecule’s next jump will be in the same channel; it equals $1-p_{sz}$ for the straight and $1-p_{zs}$ for the zigzag channel. χ is the correlation coefficient, given by $(p_{ss}^s - p_{ss}^o)/(p_{ss}^s + p_{ss}^o)$ for the straight and $(p_{zz}^s - p_{zz}^o)/(p_{zz}^s + p_{zz}^o)$ for the zigzag channel; the notation is the same we adopted in Refs. 15 and 16: p_{ss}^s is the probability of moving twice in the same direction along the straight channel, p_{ss}^o in opposite directions, and the same for the zigzag channel. The effective MSDs for one N -step sequence along x or y , to be inserted in Eq. (1), are obtained by multiplying the corresponding $\langle n^2 \rangle_{\text{seq}}$ times $(a/2)^2$ or $(b/2)^2$, respectively (a and b are the lattice constants). Finally, the MSD along z takes the form

$$\langle z^2(t) \rangle = (c/2)^2 \frac{1}{2 - (p_{sz} + p_{zs})} N_{\text{seq}}^{zc}(t). \quad (4)$$

To sum up, in the basic RW model the possible correlation involving subsequent hops between two adjacent intersections is not considered at all: this crude assumption would give good results only if the hops were completely uncorrelated. In the two-step extension the correlation between any two subsequent hops is fully accounted for: this is the strongest type of correlation present, and it can be considered as a first-order correction to the RW model. However, as shown below, the inclusion of higher-order corrections is necessary in order to properly reproduce the diffusive behavior of almost all the molecules considered in this article. While this could have been achieved by building three-, four- etc.- step correlated models analogous to the two-step one, the work of Jousse *et al.*²¹ and the results presented in this article show that this further effort is not necessary, at least for small diatomic and triatomic molecules at infinite dilution. Indeed in these cases the possible effects due to correlation extending for more than two steps can efficiently be determined by manipulating the data arising from a one-step and two-step analysis of the MD trajectory. The differences between the diffusion coefficients calculated according to the various models and those obtained through the standard MD method highlight very clearly the extent and the nature of the correlation effects influencing the diffusion in silicalite.

III. DISCUSSION

The diffusion coefficients obtained from the MSDs values computed according to the three statistical models and to the standard MD method are reported in Table I. The Einstein equation $\langle x^2(t) \rangle = 2D_x t$ has been applied to the three cartesian components, and $D = 1/3(D_x + D_y + D_z)$.

Despite the better performances of the two-step model compared to the simple RW in all cases, Table I shows that both models overestimate the diffusion coefficients, compared to the values computed directly⁸ from the MD trajectory. We highlighted¹⁶ that negative correlation effects are

TABLE I. Diffusion coefficients calculated from: (a) the standard MD method; (b) the random-walk model; (c) the two-step model; (d) the N -step model.

	D	D_x	D_y	D_z
Cl₂				
(a)	0.67±0.09	0.73±0.14	1.1±0.3	0.14±0.04
(b)	0.87	1.15	1.2	0.27
(c)	0.70	0.81	1.1	0.19
(d)	0.62	0.63	1.07	0.15
Br₂				
(a)	0.22±0.04	0.22±0.06	0.37±0.09	0.062±0.02
(b)	0.44	0.70	0.49	0.13
(c)	0.30	0.39	0.42	0.083
(d)	0.24	0.27	0.39	0.061
I₂				
(a)	0.11±0.01	0.095±0.02	0.22±0.04	0.027±0.006
(b)	0.327	0.337	0.55	0.0955
(c)	0.155	0.16	0.261	0.0452
(d)	0.103	0.10	0.18	0.029
CO₂				
(a)	0.66±0.11	0.58±0.12	1.25±0.3	0.16±0.03
(b)	1.17	1.64	1.5	0.36
(c)	0.85	0.99	1.3	0.24
(d)	0.69	0.70	1.20	0.18
CS₂				
(a)	0.28±0.05	0.19±0.04	0.59±0.15	0.06±0.01
(b)	0.656	0.84	0.92	0.20
(c)	0.386	0.44	0.62	0.10
(d)	0.289	0.31	0.49	0.066

responsible for such deviations: when the probability that the molecule performs two jumps in opposite directions along a channel is higher than the probability of moving twice in the same direction the real²³ diffusion coefficient is decreased as compared to what would be expected in the absence of correlation. From the data in Table I, it is clear that the predictions of the random-walk model are much more affected by these effects. However, the two-step model also slightly overestimates the D value in all cases: this is due to negative correlations spanning more than two jumps, because the two-step model properly discriminates between the double jumps in the same and opposite directions, i.e., it accurately includes first-order correlations. By observing the monodimensional diffusion coefficients, it appears that negative correlations (both first and higher order) are mostly present in the sinusoidal channel: $D_x^{\text{RW}} > D_x^{\text{two-step}} > D_x^{\text{MD}}$ for all considered molecules; on the other hand, in the straight channel the D_y values show the presence of considerable first-order negative correlations only for the longest species studied, iodine and carbon disulfide, while no higher order correlations are seemingly present ($D_y^{\text{RW}} > D_y^{\text{two-step}} \approx D_y^{\text{MD}}$). The trend of the D_z values is similar to that in the zigzag channel as the motion along z , proceeding through specific sequences of jumps along x and y , should be influenced by correlation effects in both channels in a more complex way.

The N -step model for the most part removes the errors in the D_x and D_z values predicted by the two-step model (due to high-order negative correlations in the zigzag channel) and the resulting agreement with the total D is very good for

TABLE II. Correlation coefficients in the straight and zigzag channels.

	χ (zigzag channel)	χ (straight channel)
Cl ₂	-0.47	-0.12
Br ₂	-0.60	-0.13
I ₂	-0.76	-0.63
CO ₂	-0.60	-0.20
CS ₂	-0.63	-0.42

all molecules. Only for chlorine the predictions of two-step and N -step models are comparable, showing that in this case almost all correlation is lost after two single jumps, i.e., the ‘‘correlation length’’ for Cl₂ is about 20 Å, while it is longer for the other molecules. It is rather surprising that CO₂, which has about the same length and mass even lower than Cl₂, shows a significantly correlated motion, i.e., first- and higher-order correlations. This proves that the presence and extent of correlation effects are not easily predictable, for instance, on the basis of pure molecular and geometrical parameters.

It is interesting to consider the values of the correlation coefficients for the straight and zigzag channels reported in Table II; they are always negative, confirming that the molecules tend to jump backward compared to uncorrelated motion. The negative correlation is more marked in the zigzag channel and it increases going to longer molecules. This agrees with the observations made on the basis of the diffusion coefficients alone; indeed, the χ value is only influenced by first-order correlations which are significant for Cl₂ also ($D^{\text{RW}} > D^{\text{MD}}$). In order to quantitatively capture the extent of higher-order correlations a different expression for the correlation coefficient would be needed. In any case, the χ values are influenced only by the tendency to go forward or backward within the same channel. Another kind of correlation concerns the tendency to remain in a channel or to turn in a different channel; the ‘‘randomization ratios’’ we introduced in Ref. 16 precisely depend on these effects. According to the N -step model, Eq. (1), the χ value affects the diffusion coefficients only because a greater (more positive) χ should lead to a greater number of jumps *per sequence*, but the total number of sequences is not influenced by it. On the other hand, the randomization ratios influence both terms in Eq. (1) because when a molecule tends to switch channel often the number of jumps per sequence decreases but the number of sequences increases. Then it is a rather difficult task to embed all these effects into a single correlation coefficient that could be quantitatively related to the diffusion coefficients, but some qualitative details of the diffusion mechanism can be separately worked out from the two types of correlation coefficients mentioned above.

An interesting capability of the N -step method is to allow the obtaining of new details concerning the diffusive motion on the basis of the two critical factors determining the MSD [Eq. (1)]: the MSD *per sequence* [related to the average number of jumps in each sequence, Eq. (3)], and the number of sequences in each channel, Eq. (2); they are reported in Table III. Note that the N_{seq} number depends linearly on the observation time: the factor in square brackets in

TABLE III. Parameters of the N -step model.

	$N_{\text{seq}}^{\text{sc}} = N_{\text{seq}}^{\text{zc}}$ ($t = 1$ ns)	$\langle n^2 \rangle_{\text{seq}}$ (zigzag channel)	$\langle n^2 \rangle_{\text{seq}}$ (straight channel)
Cl ₂	8.65	1.45	2.50
Br ₂	3.77	1.45	2.08
I ₂	2.015	1.03	1.81
CO ₂	10.9	1.28	2.23
CS ₂	4.45	1.38	2.23

Eq. (2) is approximately constant, i.e., the ratio between the parameters appearing in this factor does not change with the observation time. We put $t = 1$ ns in Eq. (2) to get an estimate of N_{seq} suitable for comparison, but the exact observation time is not relevant for this purpose.

The adimensional MSDs per sequence, $\langle n^2 \rangle_{\text{seq}}$, represent the square of the effective number of jumps in the same direction (i.e., taking into account forward and backward displacements). They vary between one and two showing that during a sequence of several consecutive jumps in the same channel a molecule normally does not depart by more than two intersections from the original one. This is connected to the high probability of two steps in opposite directions (negative correlation) already pointed out in our preceding articles.^{15,16} This effect is more marked in the sinusoidal channel and ultimately leads to the typical relation $D_x < D_y$.

The lightest molecule CO₂ carries out the largest number of sequences in the observation time but the distance covered in each sequence is shorter than for chlorine, probably on the basis of the weaker negative correlations shown by Cl₂ (see above). This results in very similar diffusion coefficients for the two species. The diffusion constants of bromine and carbon disulfide are also similar and in this case also the N -step parameters do not show marked differences. Iodine shows the lowest number of sequences in 1 ns, due to its large mass, and the MSD per sequence is also the lowest, due both to negative correlations and to kinetic reasons. In general the data of Table III show that the MSD per sequence does not vary very much for these molecules while the number of sequences carried out appears to be more crucial in determining the diffusion coefficient. In other words, for these systems correlation effects connected to the motion within the same channel, represented by χ , are likely to be less critical than correlations related to the motion between the two channel systems (embodied for example in the randomization ratios) which directly affect the total number of sequences.

IV. CONCLUSIONS

In this article we attempted to show that important data regarding the diffusion mechanism in silicalite, as well as the nature and extent of correlation effects, can be obtained by comparing the predictions of different statistical models covering different levels of correlation. The basic analysis of the molecular trajectory generated by MD provides the fundamental properties (probabilities and time lengths) of single- and double-step events. These data are enough to provide the diffusion coefficients predicted: (a) in absence of correlation;

(b) with first-order correlation included; (c) with N -step correlation included. Comparing these values between themselves and with the reference diffusion constants directly calculated by the trajectory (i.e., without any underlying statistical diffusion model), one is able to discern what kind and order of correlation affects the molecular propagation of different species. An *a priori* estimate of these data is not easy to get, or may lead to wrong guesses. For example, the different nature of correlation effects observed for Cl_2 and CO_2 molecules is hardly predictable through their physical and geometrical parameters alone. It would be highly desirable to embed all the characteristics of correlation effects into a representative single coefficient, depending on the single- and two-step data. However, we showed that it should be possible to combine the information stemming from different correlation coefficients to gain further details on the correlation effects.

Another interesting topic would be to further investigate the link between the two-step and N -step models. In all examined cases the extension to the N -step model, essentially based on the two-step parameters, resulted very appropriately. Further work is needed to test this interdependence for more complex sorbed molecules in silicalite, i.e., systems showing different adsorption and diffusion patterns.

ACKNOWLEDGMENTS

This research was supported by MURST, University of Sassari, and Consiglio Nazionale delle Ricerche.

¹J. Kärger, J. Phys. Chem. **95**, 5558 (1991).

²C. Saravanan, F. Jousse, and S. M. Auerbach, Phys. Rev. Lett. **80**, 5754 (1998).

³S. Y. Bhide and S. Yashonath, J. Chem. Phys. **111**, 1658 (1999).

⁴R. Q. Snurr, A. T. Bell, and D. N. Theodorou, J. Phys. Chem. **98**, 11948 (1994).

⁵T. Mosell, G. Schimpf, and J. Brickmann, J. Phys. Chem. **100**, 4582 (1996).

⁶S. M. Auerbach, N. J. Henson, A. K. Cheetham, and H. I. Metiu, J. Phys. Chem. **99**, 10600 (1995).

⁷S. M. Auerbach and H. I. Metiu, J. Chem. Phys. **105**, 3753 (1996).

⁸M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

⁹P. Demontis and G. B. Suffritti, Chem. Rev. **97**, 2845 (1997).

¹⁰E. J. Maginn, A. T. Bell, and D. N. Theodorou, J. Phys. Chem. **100**, 7155 (1996).

¹¹N.-K. Bär, J. Kärger, H. Pfeifer, H. Schäfer, and W. Schmitz, Microporous Mater. **22**, 289 (1998).

¹²F. Jousse, S. M. Auerbach, and D. P. Vercauteren, J. Phys. Chem. B **102**, 6507 (1998).

¹³C. Saravanan and S. M. Auerbach, J. Chem. Phys. **107**, 8132 (1997).

¹⁴F. Jousse, L. Leherter, and D. P. Vercauteren, J. Phys. Chem. B **101**, 4717 (1997).

¹⁵J. Kärger, P. Demontis, G. B. Suffritti, and A. Tilocca, J. Chem. Phys. **110**, 1163 (1999).

¹⁶P. Demontis, J. Kärger, G. B. Suffritti, and A. Tilocca, Phys. Chem. Chem. Phys. **2**, 1455 (2000).

¹⁷S. D. Pickett, A. K. Nowak, J. M. Thomas, B. K. Peterson, J. F. P. Swift, A. Cheetham, C. J. J. den Ouden, B. Smit, and M. F. M. Post, J. Phys. Chem. **94**, 1233 (1990).

¹⁸R. L. June, A. T. Bell, and D. N. Theodorou, J. Phys. Chem. **94**, 8232 (1990).

¹⁹P. Demontis, G. B. Suffritti, E. S. Fois, and S. Quartieri, J. Phys. Chem. **94**, 4329 (1990).

²⁰C. J. J. Den Ouden, B. Smit, A. F. H. Wielers, R. A. Jackson, and A. K. Nowak, Mol. Simul. **4**, 121 (1989).

²¹F. Jousse, S. M. Auerbach, and D. P. Vercauteren, J. Chem. Phys. **112**, 1531 (2000).

²²In the N -step model the sum of the probabilities of the two-step events is 1 separately for each channel, while in our previous paper (Ref. 16) the sum of all probabilities (including events in both channels) is 1, so that they have to be recalculated according to the formulation of Jousse *et al.*

²³The diffusion coefficients determined by MD (Ref. 8) (with the present force field) and by experimental methods such as (pulsed field gradient) PFGNMR are in good agreement in the case of ethane (Refs. 15 and 24) and CO_2 (Ref. 16), for which experimental data are available. Therefore, we assume that the MD data are a reasonable approximation to the real data for short, linear molecules, even for systems where no experimental data are available.

²⁴P. Demontis, G. B. Suffritti, and A. Tilocca, J. Chem. Phys. **105**, 5586 (1996).