Bond fluctuation method for a polymer undergoing gel electrophoresis

Ryuzo Azuma and Hajime Takayama

Institute for Solid State Physics, The University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8666, Japan
(Received 11 August 1998)

We present a computational methodology for the investigation of gel electrophoresis of polyelectrolytes. We have developed the method initially to incorporate sliding motion of tight parts of a polymer pulled by an electric field into the bond fluctuation method (BFM). Such motion due to tensile force over distances much larger than the persistent length is realized by nonlocal movement of a slack monomer at either end of the tight part. The latter movement is introduced stochastically. This new BFM overcomes the well-known difficulty in the conventional BFM that polymers are trapped by gel fibers in relatively large fields. At the same time it also reproduces properly equilibrium properties of a polymer in a vanishing field limit. The new BFM thus turns out to be an efficient computational method to study gel electrophoresis in a wide range of the electric field strength. [S1063-651X(99)01401-4]

PACS number(s): 83.10.Nn, 61.25.Hq

I. INTRODUCTION

Gel electrophoresis is a method which separates polyelectrolytes such as DNA according to their length. This technique is an application of the phenomenon in which polyelectrolytes exhibit different migration velocities in gel when they are pulled by an external electric field. Although in recent years many sophisticated methods such as pulsed-field techniques have been developed [1,2], its underlying physics, even merely on steady field techniques, are not thoroughly understood.

The concept of reptation, which was proposed by de Gennes [3], has been incorporated into theoretical analysis of gel electrophoresis. For example, the biased reptation model has succeeded in explaining an empirical law of steady field experiments in the small field limit [4,5], i.e., DNA mobility is proportional to the reciprocal of its length [6]. However, such predictions by the reptation theory are only on properties associated with dynamics of averaged conformations of DNA.

It is now well known that DNA dynamics in electrophoresis is more complicated and interesting. The molecular-dynamics (MD) simulation by Deutsch [7] on a freely jointed chain in a two-dimensional space with obstacles substituted for gel fibers first demonstrated that the chain migrates through obstacles taking extended and collapsed conformations alternatively. Since this oscillatory behavior appears as a steady state [8], the averaged conformation which the reptation model predicts is not stable. Also experimentally, since fluorescent microscopy has been invented and improved, evidence showing this instability has increased as well, and now details of DNA motion itself are of current interest in the study of gel electrophoresis [9–12]. Among them the inch-worm-like motion may be the most typical example which is observed for large DNA [13].

In studying polymer dynamics, numerical simulations have played important roles. One of them is the MD method, whose example was already mentioned above. Another powerful simulation is the bond fluctuation method (BFM), which is a Monte Carlo method utilizing a description of a polymer(s) on a lattice [14]. Its efficiency in examining gel electrophoresis in a small field was already demonstrated, but at the same time, its difficulty when applied to the case in a large field was pointed out [15]. The difficulty is that it takes huge MC steps (MCS's) for polymers once hanging on obstacles to get rid of them. As a result mobility in a large field becomes significantly smaller than in a relatively small field. This is an artifact of the numerical method since in actual experiments the mobility is observed to increase monotonically with increasing field [16]. A reason for this difficulty may be that tensile force between monomers, which are of fundamental importance when each part of the chain is in nearly straight conformations, is not taken into account in the conventional BFM (c-BFM).

The purpose of the present paper is to report a BFM (n-BFM) which we have developed to overcome the above-mentioned difficulty by introducing to the c-BFM new stochastic processes which simulate sliding motions caused by tensile force. The method turns out to be able to reproduce qualitative aspects of gel electrophoresis phenomena in a wide range of the field. In large fields polymer motion is free from trapping by obstacles and mobility monotonically increases (and tends to saturate) with increasing field as we have expected. In small fields, on the other hand, the results simulated by the method coincide with those simulated by the c-BFM if the time unit is properly scaled. This means that, in this field range, the stochastic process newly introduced here contributes to the entropy effect due to polymer conformations similar to what the c-BFM process does. Furthermore, in a certain limited range of field we have observed extended and contracted conformations of a polymer very frequently, though not quasiperiodically. The n-BFM is considered very efficient, complementary to the MD method, in studying gel electrophoresis.

In the next section we describe the n-BFM in detail. One further detail in the algorithm is explained in Appendix A. The results simulated by the n-BFM are presented and compared with those obtained by the c-BFM in Sec. III, and the final section is devoted to concluding remarks of the present work.
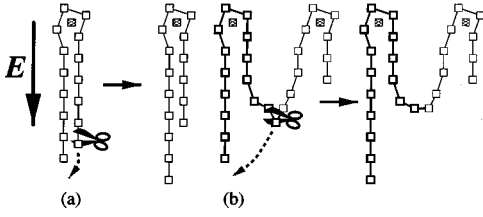
FIG. 1. Schematic illustration of (a) the U-shaped and (b) M-shaped conformations pulled down by the field.

II. METHOD

In the present work we consider a $d=2$ square lattice version of the bond fluctuation method (BFM) [14] to simulate gel electrophoresis. In this method each monomer is represented by a unit cell (four lattice sites) and each bond has a variable length l but with the restriction $2 \leq l \leq \sqrt{13}$ (see Fig. 11 in Appendix A). A monomer cannot fill a site which is already occupied by other monomers or by obstacles. The latter are substituted for gel fibers and each of them consists also of a unit cell. In the present work they are distributed periodically in both x and y directions with period of a lattice units. A uniform, steady electric field is applied in the (1,1) direction. We use dimensionless electric field strength $E = q\mathcal{E}e/k_BT$, where q, $\mathcal{E}$, e, k_B and T denote charge of a monomer, bare electric field strength, lattice constant ($\sim$ persistent length, and is set to unity in the present work), Boltzmann constant, and temperature, respectively. The MC processes in the BFM consist of local random walks (of unit length) of each monomer in a potential due to the electric field. The excluded volume effect is satisfied and no bond crossing occurs in this BFM, which we call the conventional BFM (c-BFM). Actual simulations are done on square lattices of sizes $3M \times 3M$ with periodic boundary conditions. Here M is the number of monomers in the chain.

Although the c-BFM described above has proven its efficiency in examining gel electrophoresis in smaller fields, it has serious difficulty in larger fields [15]. The latter is easily understood from the inspection of a case shown in Fig. 1(a), where we show a schematic picture of the U-shaped conformation of a chain hooked by an obstacle and pulled by a large downward field. Within the c-BFM the chain can escape from the trap only by the following process: a relatively slack part of the chain, created around the end segment, climbs up sequentially against a potential slope of the field up to the position of the obstacle. The probability for this process to occur is smaller the larger the field is. For ex-

ample, the mobility μ of the $M=200$ chain simulated by the c-BFM increases with increasing E up to $E \simeq 0.01$, but it starts to decrease and tends to vanish when E is increased further. Here the mobility μ is simply evaluated by

$$\mu = \frac{\langle X_G(t_f) - X_G(t_i) \rangle}{t_f - t_i} \Big/ E \qquad (1)$$

in our units, where $\langle \cdots \rangle$ indicates the average over the MC runs. In Eq. (1) X_G is the displacement of the center of mass in the field direction. The starting time t_i of observation is chosen to eliminate the influence of an initial conformation, and t_f is chosen by the condition that $\langle X_G(t_f) - X_G(t_i) \rangle$ is larger than $c_\mu a$ or $c_\mu R_I$ with $c_\mu \gtrsim 10$, where R_I is the radius of gyration of the chain.

In actual experiments DNA can slide off an obstacle even if it is temporarily trapped by the latter. For example, Volkmuth and Austin fabricated a quasi-two-dimensional microlithographic array of posts and looked at a sequence of moving images of 100 kb DNA in a steady field of 1.0 V/cm (corresponding to $E \simeq 0.005$)[17]. They observed that the DNA hooks one of the posts, is extended to nearly its full contour length, and then slides from the post. This indicates that tensile force plays an important role in such a sliding process because the chain is fully extended in the process.

It is then natural to introduce to the c-BFM the following nonlocal processes which simulate realistic sliding motions of DNA due to tensile force. In the case of the U-shaped conformation of Fig. 1(a), we simply remove the end monomer of the shorter arm and join it to the other end of the chain. For the M-shaped conformation shown in Fig. 1(b), we remove one of the monomers in a dangling part at the center to the end of the longer arm.

Starting from the above intuitive idea, we have constructed an algorithm in which nonlocal dynamics of monomers is systematically and stochastically introduced. For this purpose we first define monomers, on which such nonlocal processes are tried. They should be in a slack part of the chain, through which tensile force cannot transmit. We call them slack monomers (s monomers). The s monomers defined consist of monomers at both ends of a chain and of monomers whose nearest neighboring monomers are not separated from each other by distance larger than or equal to 4. Parts of the chain between neighboring s monomers are regarded as tight parts, through which tensile force transmits. They are considered to slide either partially or as a whole depending on nonlocal movement of s monomers which we next introduce.

Our guiding principle to specify a procedure of non-local movement of s monomers is to choose such stochastic processes that any s monomer moves so as to fulfill the detailed balance condition in equilibrium. This restriction, however, does not specify a procedure uniquely. Among possible procedures we adopt the following. (i) Choose one s monomer [monomer 0 in Figs. 2(a) and 2(b) shown for examples]. (ii) Count the number of pairs n, on which we can try to move the chosen s monomer, and which are in a region between its neighboring s monomers including ends of the chain (pairs 12, 23, 34, and 45 for case 2(a) ["end" for case 2(b)] between monomers 0 and 5 ("end"), and $1'2'$, $2'3'$, $3'4'$, and $4'5'$ between monomers 0 and $5'$, and so $n=8$). (iii) Choose
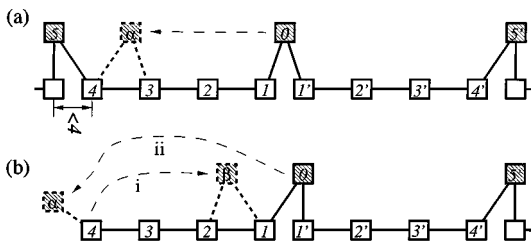

FIG. 2. Examples of s monomers and their movements. (a) The case in which no end monomer is involved. (b) The case in which an end monomer is involved.
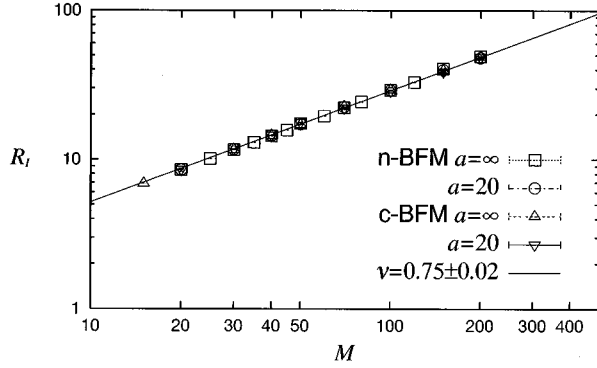
FIG. 3. Dependence of radius of gyration R_I on length M.

randomly one of these pairs [34 (''end'')] with the probability $1/n$. (iv) Choose randomly one of the allowed conformations putting a monomer between the chosen pair [$3\alpha4$ (4α)] with the probability $1/W$. By construction the moved monomer is an s monomer. (v) Accept the movement of the s monomer [from s monomer 0 to s monomer α] according to the weight $w(\Delta X)$ defined by

$$w(\Delta X) = \frac{e^{E\Delta X}}{e^{-E\Delta X} + e^{E\Delta X}}, \qquad (2)$$

where ΔX is the displacement of the s monomer in the (positive) direction of the field.

In (iv) above, the value of W is chosen to be greater than $\max\{W_1, \ldots, W_n\}$, where W_i is the number of the allowed conformations putting a monomer to the ith pair. In the present work we fix $W = 23$, which is the possible maximum value of W_i when the ith pair is one of the ''ends.'' Then the probability of the movement of the s monomer to one specified conformation is given by $1/nW$ multiplied by a weight $w(\Delta X)$, while that of its reversed process is $1/nW$ multiplied by weight $w(-\Delta X)$. This guarantees the detailed balance condition. It is noted here that the acceptance ratio of one nonlocal movement without specifying the moved conformation at all, r_{move}, is given by $r_{\text{move}} \simeq \overline{W_i}/2W$ in the limit $E\Delta X \to 0$, where $\overline{W_i}$ is the average of W_i's including the further restriction described below.

There remain, however, two problems which require further considerations. One is concerned with the detailed balance condition. By the procedure described above it may
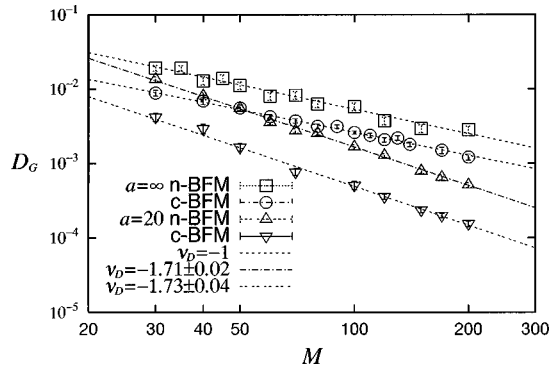

FIG. 4. D_G-M plots for $a = 20, \infty$ simulated by the two BFMs. ν_D denotes the exponent of the power-law dependences.

happen that the number of s monomers changes depending on the positions of the monomers around the chosen pair [in case Fig. 2(a), besides the α monomer, monomers 3 or 4 may become an s monomer]. If this happens, the number n for the moved s monomer also changes after the movement (except for the case that the movement involves an ''end'' s monomer). Then, by the above procedure the probability of the reverse process violates the detailed balance condition. We get rid of this problem simply by accepting only processes by which the number n for the moved s monomer is conserved.

The other problem remaining is how to exclude possible bond crossing which may occur when an s monomer is moved to an end of the chain. We are faced with the same problem when we try to construct an initial conformation of a chain by the self-avoiding random walk. We have developed an algorithm which recognizes whether bond crossing occurs or not by referring to only local data. It is explained in Appendix A.

The new BFM (n-BFM) we propose is a combined process of nonlocal movements of s monomers by the above procedure and the c-BFM process: each odd (even) MC step of the c-BFM is followed by trials of nonlocal movements of odd (even) numbered s monomers. [For a local movement of monomers in the c-BFM in the present work, we have adopted the same weight $w(\Delta X)$ with $\Delta X = \pm 1/\sqrt{2}$ after choosing one of the four neighboring sites to move with equal probability.]

III. RESULTS

Let us begin with a comparison of equilibrium properties in $E = 0$ simulated by both the c-BFM and the n-BFM. As for a representative of static equilibrium properties, we show four sets of data of radius of gyrations R_I in Fig. 3 in cases with obstacles ($a = 20$) and without obstacles ($a = \infty$) obtained by the two BFM's. They all coincide with each other and fit well with the power law $R_I \propto M^\nu$ with $\nu = 0.75 \pm 0.02$ which is consistent with the previous results [18]. There is no significant difference between the cases $a = 20$ and $a = \infty$ in R_I-M dependences. This is because the concentration of obstacles is too low to change the value of exponent ν [2].

In Fig. 4 the diffusion constants D_G are plotted against M to compare behaviors of fluctuation simulated by the two BFM's. Here D_G is evaluated simply by

$$D_G = \frac{(\Delta R_G)^2}{2(t_f - t_i)} (\Delta R_G)^2 = \langle [\mathbf{R}_G(t_f) - \mathbf{R}_G(t_i)]^2 \rangle, \qquad (3)$$

where $\mathbf{R}_G$ is the coordinate of the center of mass of the chain. The data actually used are those in the range $R_I \lesssim \Delta R_G \leq c_D R_I$, where c_D is at least larger than 2. In the case without obstacles both BFM's yield $D_G \propto M^{-1}$ as expected, while in case $a = 20$ the same power-law dependence $D_G^{-\nu_D}$ with $\nu_D = 1.71 \pm 0.02$ (n-BFM) and 1.73 ± 0.04 (c-BFM) is obtained. However D_G by the n-BFM is larger by factor 2 ~ 3 than that by the c-BFM.

A more stringent check of the new algorithm may be whether it reproduces the fluctuation-dissipation theory or the Einstein relation $\mu/M = D_G$, where μ is the mobility in
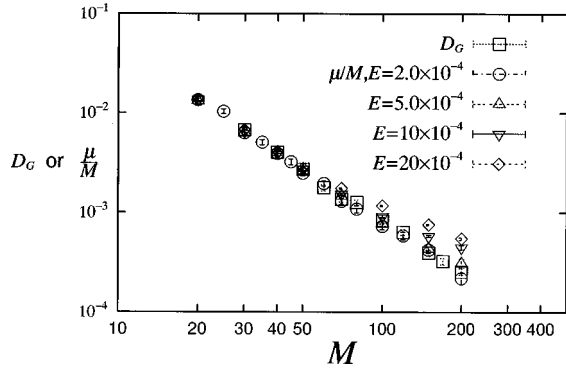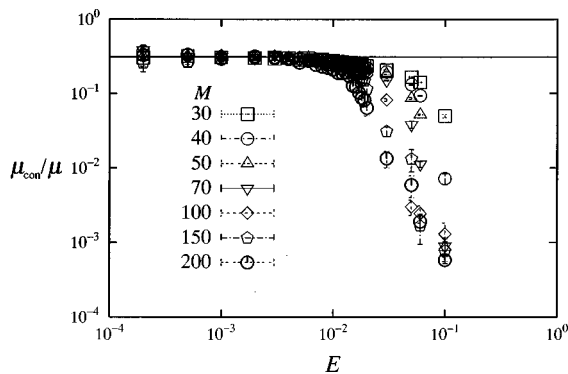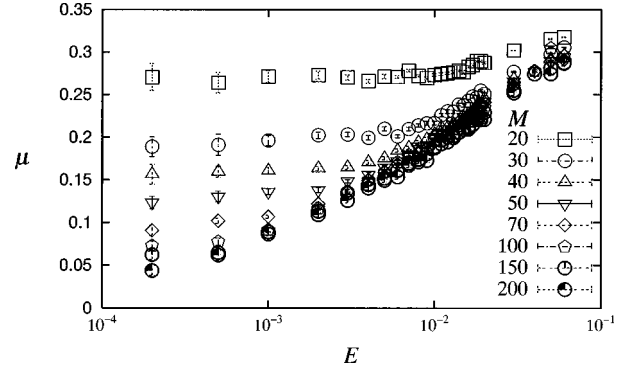
FIG. 5. Diffusion constant D_G and μ/M plotted against M.

FIG. 7. Plots of μ vs ln$E$.

the limit $E \to 0$. We plot μ/M for various E and D_G against M in Fig. 5. Clearly we see that the Einstein relation holds well. The results described so far indicate that our new algorithm reproduces equilibrium properties of a polymer quite satisfactorily.

Let us next examine stationary properties in finite E simulated. In order to compare those obtained by the two BFM's, we show the ratio $\mu_{\rm con}/\mu$ for various M with $a = 20$ plotted against E in Fig. 6. Here $\mu_{\rm con}$ and μ are mobilities simulated by the c-BFM and the n-BFM, respectively. In small fields $E \lesssim 5.0 \times 10^{-3}$ the ratio is almost constant and is also independent of M. The constant value is equal to the ratio of the diffusion constants D_G of the two BFM's mentioned above.

The ratio $\mu_{\rm con}/\mu$ deviates from the constant value when E exceeds a certain value $E_{\rm th}$ which depends on M. Its deviation is more significant for chains with larger M, and the ratio tends to vanish at E much larger than $E_{\rm th}$. This result simply reflects the difficulty of the c-BFM described in the preceding section and is considered to have nothing to do with the n-BFM. Actually there seems no crossoverlike behavior around $E_{\rm th}$ in μ obtained by the n-BFM as shown in Fig. 7. One sees in the figure that μ of each M increases monotonically with increasing E, and tends to saturate at largest $E(\sim 0.1)$ we have simulated. Thus the data in Figs. 6 and 7 tell us that we have in fact overcome the difficulty of the c-BFM by means of the n-BFM we have proposed.
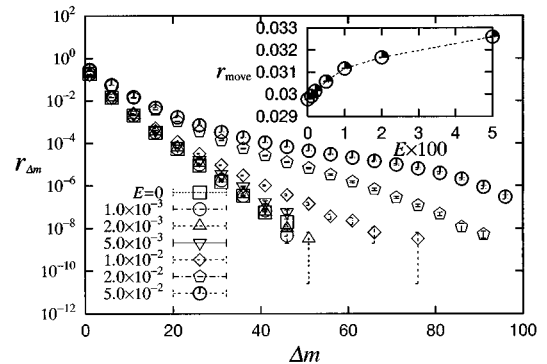
In order to get further insights into the n-BFM, which gives rise to quite satisfactory results demonstrated so far, we have examined some details of the s-monomer movement in a system with $M = 100$ and $a = 20$. Interestingly, the accep-

tance ratio of the s monomer movement, $r_{\rm move}$, defined in the preceding section is rather small ($\sim 3\%$) and does not depend sensitively on E in the range examined as shown in the inset of Fig. 8. On the other hand, the average number of the s monomers, $\bar{M}_s$, is rather large: $\bar{M}_s \simeq 28$ at $E = 0$ and $\bar{M}_s \simeq 24$ at $E = 0.05$.

These results are interpreted as follows. In the field range investigated here ($E \lesssim 0.05$), the entropy effect dominates to determine conformations of a chain. The latter is therefore in a coiled conformation on average. The s monomer movement in this situation is regarded as a local fluctuation of the conformation due to the entropy effect rather than due to the tensile force. Although $r_{\rm move}$ is rather small, the s-monomer movements contribute to displacement of the center of mass of the chain which is comparable in magnitude with that by one MCS of the c-BFM process. This explains why D_G and μ by the n-BFM are larger than the corresponding quantities by the c-BFM (note that one mcs in the n-BFM consists of one MCS of the c-BFM process and trials of nonlocal movement on half of the s monomers). Such efficiency of the s monomer movement to D_G and μ may be attributed to the fact that the s monomer moves along the chain direction which is hardly realized by the c-BFM process.

The main part of Fig. 8 is the histogram of the relative ratio of the s-monomer movements accepted, $r_{\Delta m}$, against the distance (along the chain), Δm, they moved. Notice that the movement with $\Delta m \sim M/2$ occurs even at $E = 0$, though

FIG. 6. The mobility ratio $\mu_{\rm con}/\mu$ plotted against E, where $\mu_{\rm con}$ is obtained by the c-BFM and μ by the n-BFM.

FIG. 8. Histograms of the s-monomer movements accepted, $r_{\Delta m}$, against the distance (along the chain) Δm for the chain with $M = 100$ and $a = 20$. In the inset is shown the acceptance ratio of the s-monomer movement, $r_{\rm move}$.
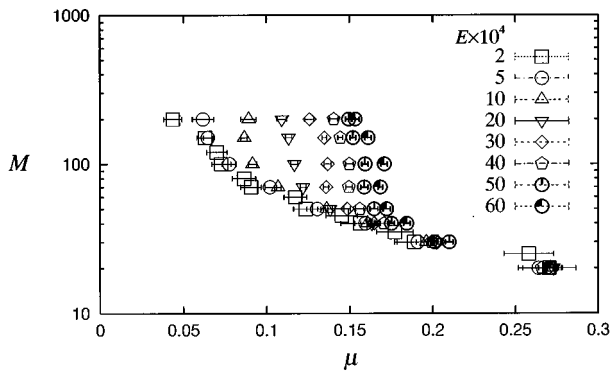
FIG. 9. Plots of $\ln M$ against μ.

$r_{\Delta m}$ is quite small (the abscissa is in the logarithmic scale). This $r_{\Delta m}$ at $E=0$ reflects how frequently, in fact quite rarely, extended conformations occur by fluctuation in the coiled conformation of the chain.

An important observation of Fig. 8 is that $r_{\Delta m}$ for $E \geq 0.02$ is much enhanced at large Δm, though the absolute magnitude is still quite small. This means that even in this range of E extended conformations occur with low probability. But once they are present, the s-monomer movements with large Δm occur with relatively high probability. This situation is just what we have first intuitively expected: the tensile force is considered to play an important role on the dynamics of the chain. The low probability of such nonlocal movements of the s monomer, which nevertheless overcomes the difficulty in the c-BFM and reproduces smooth, monotonic E dependence of the mobility, justifies our introduction of the s monomer movements as a stochastic process.

In Fig. 9 we show the μ versus $\ln M$ plot using the same data shown in Fig. 7. It can be seen that μ for each E decreases monotonically as M increases, and that they are strongly dependent on M when E is small enough. As E increases, however, it gradually loses M dependence from larger values of M, and finally μ becomes almost independent of M when E is quite strong ($E \sim 0.1$). This behavior of μ is qualitatively consistent with experimental results [16].

Finally we note that chain conformation dynamics simulated by the present n-BFM is actually complicated. Particu-
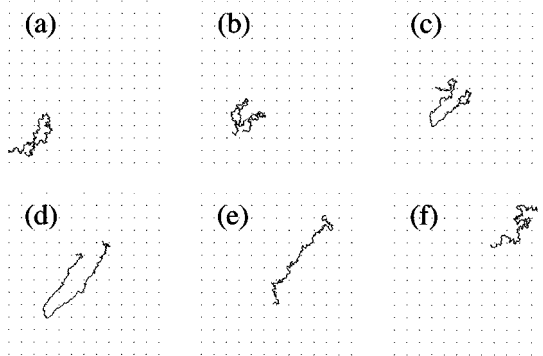

FIG. 10. Time development, (a), (b), . . . , (f), of $M=200$ chain conformations in field $\mathbf{E} \| (1,1)$ and $E = 1.02 \times 10^{-2}$. The snapshots at every 4.0×10^4 MCS are shown. Grid points represent obstacles ($a=30$).
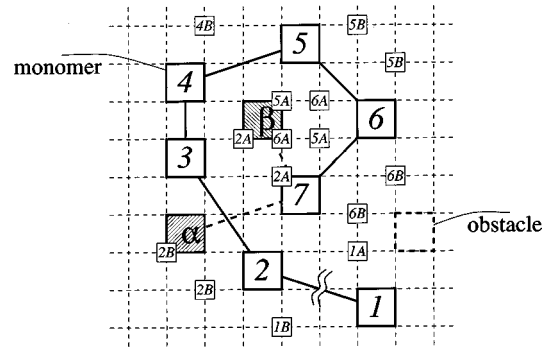

FIG. 11. An example of conformations of an $M=7$ BFM chain on a square lattice. The unit square with thick broken lines represents an obstacle. Small squares on sites denote the local function $\Phi(n)$.

larly in a certain limited range of E (~ 0.01), sequences of contraction and extension are observed. One such sequence is shown in Fig. 10, which shows snapshots of conformation of an $M=200$ chain in every 4×10^4 MCS in a certain MC run. A chain moving in a relatively collapsed form (a) is trapped by a certain obstacle (b), deforms to a V- or U-shaped conformation (c) and (d), slides off the obstacle (e), and then tends to form a collapsed conformation again (f). In contrast to experimental observation [13], however, quasiperiodical alternation between contracted and extended conformations has not yet been observed, or periods between the two conformations are rather random.

IV. CONCLUDING REMARKS

By introducing stochastic, nonlocal movements of slack parts of the polymer (s monomers) into the conventional bond fluctuation method (c-BFM), we have constructed a new BFM (n-BFM) algorithm which overcomes the difficulty of the c-BFM applied to gel electrophoresis in a relatively large field, namely, polymers once hooked by gel fibers become unable to get rid of them. In smaller fields, on the other hand, the new stochastic process gives rise to conformational change of the polymer which is interpreted as the ordinary entropy effect. The present n-BFM thus turns out to be able to reproduce qualitative aspects of gel electrophoresis phenomena in a wide range of the field.

The n-BFM is considered more effective for a denser gel (a smaller a). Actually the preliminary analysis of the mobility in case $a=12$ yields $\mu_{\text{con}}/\mu \simeq 0.25$ at small E, which is smaller than that of the case $a=20$ shown in Fig. 6. This tells that nonlocal s monomer movements make it easier for parts of the polymer to escape obstacles, and so even the entropic conformation change in smaller fields is fastened as compared with the c-BFM.

There are many problems to be explored further. Among them are an improvement of algorithms to increase the acceptance ratio of the s monomer movements, r_{move}, without violating the detailed balance, and quantitative comparisons with experimental observations. For the latter purpose we have to adopt model systems with more realistic distribution and/or shape of obstacles, and to extend the algorithm to $d = 3$ system. These problems are now under investigations.

APPENDIX A: REJECTION PROCEDURE FOR BOND CROSSING

Here we explain a method of rejecting a nonlocal movement of an s-monomer to one end of a chain which associates with bond crossing. For this purpose we use, as a simple example, an $M=7$ BFM chain drawn in Fig. 11. In addition to monomers and connecting bonds, we introduce a local function $\Phi(n)$ represented by small squares on each site n. This function contains information on the position of each site n relative to nearby bonds which are now specified as vectors $\overrightarrow{i,i+1}$. When $\Phi(n)$ takes a value ''iA'' (''iB''), it means that site n is near the bond $\overrightarrow{i,i+1}$ (both distances from monomers i and $i+1$ are less than 4) and is in its right-(left-) hand side. Since site n can be near other bonds, $\Phi(n)$ is, in general, multivalued. If, on the other hand, site n is far from any bonds $\Phi(n)=$ ''null.''

By making use of the local function $\Phi(n)$ thus defined, we can check bond crossing as follows. To be simple, let us consider the case removing end monomer 1 and putting it to the other end monomer 7, as shown in Fig. 11. If monomer 7 touches a site with $\Phi=$ ''$2A$'', the moved monomer should not touch sites with $\Phi=$ ''$2B$''. Monomer α is such a case where bonds $\overrightarrow{2,3}$ and $\overrightarrow{7,\alpha}$ cross and therefore is rejected. Other cases such as monomer β are all accepted from the present criterion on bond crossing.

[1] B. Norden, C. Elvingston, M. Jonsson, and B. Åkerman, Q. Rev. Biophys. **24**, 103 (1991).
[2] B. H. Zimm and S. D. Levene, Q. Rev. Biophys. **25**, 171 (1992).
[3] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).
[4] O. J. Lumpkin, P. Déjardin, and B. H. Zimm, Biopolymers **24**, 1573 (1985).
[5] G. W. Slater and J. Noolandi, Biopolymers **25**, 431 (1986).
[6] N. C. Stellwagen, Biopolymers **24**, 2243 (1985).
[7] J. M. Deutsch, Science **240**, 922 (1988).
[8] J. M. Deutsch and T. L. Madden, J. Chem. Phys. **90**, 2476 (1989).
[9] S. B. Smith, P. K. Aldridge, and J. B. Callis, Science **243**, 203 (1989).
[10] D. C. Schwartz and M. Koval, Nature (London) **338**, 521 (1989).
[11] S. Gurrieri, E. Rizzarelli, D. Beach, and C. Bustamante, Biochemistry **29**, 3396 (1990).
[12] T. D. Howard and G. Holzwarth, Biophys. J. **63**, 1487 (1992).
[13] H. Oana *et al.*, Macromolecules **27**, 6061 (1994).
[14] I. Carmesin and K. Kremer, Macromolecules **21**, 2819 (1988).
[15] J. Batoulis, N. Pistoor, K. Kremer, and H. L. Frisch, Electrophoresis **10**, 442 (1989).
[16] H. Hervert and C. P. Bean, Biopolymers **26**, 727 (1987).
[17] W. D. Volkmuth and R. H. Austin, Nature (London) **358**, 600 (1992).
[18] D. S. Mckenzie, Phys. Rep. **27**, 35 (1976).