# *Predicting* equilibrium structures in freezing processes

Dieter Gottwald
*Center for Computational Materials Science and Institut für Theoretische Physik, Technische Universität Wien, Wiedner Hauptstraße 8-10, A-1040 Wien, Austria and Institut für Theoretische Physik II, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany*

Gerhard Kahl
*Center for Computational Materials Science and Institut für Theoretische Physik, Technische Universität Wien, Wiedner Hauptstraße 8-10, A-1040 Wien, Austria*

Christos N. Likos
*Institut für Theorestische Physik II, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany*

(Received 15 February 2005; accepted 11 March 2005; published online 23 May 2005)

We propose genetic algorithms as a new tool that is able to *predict* all possible solid candidate structures into which a simple fluid can freeze. In contrast to the conventional approach where the equilibrium structures of the solid phases are chosen from a preselected set of candidates, genetic algorithms perform a parameter-free, unbiased, and unrestricted search in the entire search space, i.e., among all possible candidate structures. We apply the algorithm to recalculate the zero-temperature phase diagrams of neutral star polymers and of charged microgels over a large density range. The power of genetic algorithms and their advantages over conventional approaches is demonstrated by the fact that new and unexpected equilibrium structures for the solid phases are discovered. Improvements of the algorithm that lead to a more rapid convergence are proposed and the role of various parameters of the method is critically assessed. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1901585]

## I. INTRODUCTION

The freezing behavior of simple fluids can be predicted, as a consequence of considerable progress achieved during past years, very accurately. The conventional approach to this problem proceeds along the following lines: first—relying on experience, intuition, or plausible arguments—a set of possible candidate structures for the solid phase is preselected. Then, using microscopic liquid state theories, the free energy of the fluid phase is calculated.[1,2] For the solid, a wide spectrum of methods, at varying levels of sophistication, are available for the calculation of the free energy of the above-mentioned preselected crystal structures. These methods include, with increasing numerical and conceptual effort, simple lattice sums,[3] cell models,[4–7] harmonic solid theories of various degrees of sophistication,[8,9] and modern density functional approaches.[10–14] In this way, one can calculate the free energies of the competing phases and compare the thermodynamic stability of the respective solid structures. For a given temperature and density, the free-energy calculations along with the double tangent construction provide then information on the stability of the crystalline phases and the location of possible phase boundaries between them. This approach can also be applied, albeit with considerably increased numerical effort, to more complex systems that are, for instance, characterized by an additional internal degree of freedom.[15] The inherent weak point of this approach is the preselection process, which bears the risk to simply "forget" possible equilibrium structures. In this contribution, we introduce the tool of genetic algorithms (GA) to address this problem. We will demonstrate that GAs are able to determine the structure of the solid phases via a parameter-free, unbiased, and unrestricted search strategy in the *entire* search space.

GAs were originally developed by Holland and co-workers[16] and have later been extended to different fields (such as business, economics, or engineering; see, e.g., Ref. 17), where they have meanwhile become a very attractive tool due to several reasons: no particular requirements are imposed on the function to be optimized and restrictions on the possible solutions can be included very easily. Further, GAs have turned out to be numerically robust, powerful, and well suited for problems in large and complex search spaces; they directly aim at global maxima and not only at the nearest local one; finally, their numerical implementation is rather simple. Although GAs are meanwhile widely used in different fields, they have not yet become particularly popular in physics or chemistry. To the best of our knowledge, they have been used up to now only for optimizing cluster geometries (see Ref. 18, and references therein).

In several of the optimization problems arising in our field of research, the theory of condensed matter, we have realized that GAs are a very useful, efficient, and reliable tool. In this contribution we want to demonstrate its potential, power, and merits in an application to a standard problem of condensed matter: the freezing transition. We point out the particular advantages of the GA over conventional approaches and want to encourage with this contribution condensed matter theoreticians to a more widespread use of this attractive tool.

 

The conventional approach to the freezing process as outlined above has—despite its merits—the drawback that it is not able to *predict* in an unbiased way those crystalline structures into which the fluid phase freezes: the preselection of possible candidate structures in the beginning already rules out all other structures, which are either too complex to be taken into account or which simply have not been considered at all; they will never show up in the phase diagram even though they might represent possible stable crystalline structures. The conventional access to freezing therefore always carries the risk that results are not predicted in a constructive way but might rather be biased by *a priori* assumptions. It is therefore highly desirable to search for an alternative tool, that is, able to *predict* those stable structures into which a simple (or even complex) fluid can freeze in an unbiased and unrestricted search—possibly in the entire search space—without excluding any structure from the very beginning. The absence of any "culturally induced prejudices" regarding the candidate crystal structures (which oft leads to a prejudice in favor of the cubic lattices) is of tantamount importance when dealing with effective interactions that typically arise between macromolecular aggregates, usually encountered in soft condensed matter physics. Indeed, the ultrasoft character of these potentials[19] often leads to the emergence of "exotic" crystalline structures[3] with very unusual point-group symmetry properties.[20,21] Our investigations provide convincing evidence to the fact that GAs are indeed able to fulfill these expectations, as will be demonstrated in what follows.

As mentioned above, GAs are optimization strategies; they can be viewed as an implementation of Darwin's process of evolution, using thus features of evolutionary processes as key elements to find the optimal solution for a problem. The basic quantity in a GA is an *individual* which represents a candidate solution; individuals are evaluated via a problem-specific fitness function in the sense that a better (or fitter) individual has a higher fitness value. A large number of individuals forms a *generation*. After selecting—according to their fitness values—parents in one generation, individuals of the subsequent generation are formed via simple operations (such as recombination and mutation). Iterating through several generations, retaining in each generation the individual with the highest fitness value, and taking in the end the one with the absolutely highest fitness value leads to the final solution.

In this contribution we apply the GA to the freezing problem and consider two standard systems of soft matter: neutral star polymers and charged microgels. We have focused on systems interacting by soft potentials rather than by hard ones, since the equilibrium structures of the latter are by now well studied and bear no further secrets. On the other hand, investigations of the phase behavior of soft systems have brought along, in recent times, many surprises and unexpected results. Therefore, they represent an ideal testing ground for the GA. Indeed, we are able to demonstrate the power of this algorithm by comparing our results for neutral star polymers with those presented in Ref. 3, which were obtained via a conventional approach: even though ponderous considerations in combination with very sophisticated concepts[20,21] were applied in Ref. 3 to predict the solid phases, our GA predicted with its parameter-free, unrestricted, and unbiased search algorithm new and unexpected equilibrium structures for the system.

In our application of the GA to the freezing problem, we identify a possible lattice structure at a given temperature and density as an individual; this representation is realized via a suitable translation rule, which serves to express geometrical properties in a binary language. Selecting parents of one generation via their respective fitness values (which are related to the energy to this lattice) lead via the usual cycle (selection-recombination-mutation-evaluation) to individuals of the subsequent generation. The individual with the absolutely highest fitness value, i.e., the structure with the lowest free energy, is considered—after a final application of a hill-climbing search—as the solution to the problem and thus as the equilibrium structure for a given temperature and density. Our results for the phase diagrams of soft materials are then complemented by a critical discussion on refinements of the concept and of the numerical parameters, and of the implementation of the code. We confirm that GAs can indeed be implemented very easily, they are robust and reliable, and converge fast.

We conclude this introduction by a brief summary of a recent study on the mating behavior of *Parus caeruleus*[22] which demonstrates the nature's basic strategies to guarantee survival of a species are realized in GAs: ornithologists found out that extra-pair mating of the females leads to a fitter offspring: while social monogamy reduces the reproductive success (expressed, e.g., by a higher susceptibility to diseases), mating between genetically dissimilar species leads to an offspring with a higher genetic quality. This is reflected in a higher reproductive success and increased survival chances, the offspring is genetically diversified, and has a higher fitness. We shall establish a relation between these findings and our results later in this contribution.

The rest of the paper is organized as follows: In Sec. II we introduce the basic ideas of a GA and describe how they are translated to address the freezing problem. In Sec. III we present results obtained for two typical soft matter systems. In Sec. IV we propose refinements of the GA technique and discuss the role of its numerical parameters. Finally, in Sec. V we summarize and draw our conclusions.

## II. THEORY

### A. Genetic algorithms in general

We start with a brief summary of the basic ideas of a GA and will show in the subsequent section how this scheme translates to our specific problem. In a GA an individual (genotype) represents a point in search space and thus one candidate solution (phenotype). The genotype is built up by a fixed number of genes. In our case, one individual will represent the primitive vectors that form a simple lattice in a coded form; eventually it will also include positions of further particles in a nonsimple lattice, possibly a vector for an additional internal degree of freedom, or further parameters that characterize the geometrical structure under consideration. The genes in an individual can take values out of a

suitably chosen alphabet. A large but *fixed* number of individuals forms a generation. Further, a positive-definite fitness function has to be introduced, which evaluates the individuals in the sense that a higher fitness value represents a better solution. Constraints on the individuals which stem from constraints on the phenotypes can easily be taken into account: either via a penalty function which reduces the fitness of an individual or by a suitable encoding (parametrization) of the same. The GA proceeds as follows: the initialization of the first generation is realized at random, then pairs of individuals are selected according to their fitness value. These pairs represent parents that generate individuals of the subsequent generation; several procedures for this mating (recombination) processes have been proposed in the literature.[17] In addition, mutations are performed on these individuals with a probability $p_{\mathrm{mutate}}$, a procedure necessary to avoid persistent inbreeding and allowing at the same time for reintroduction of new or lost genetic material. Within one generation we retain the individual with the highest fitness value in a list. With the individuals of the second generation we proceed as in the preceding generation. This cycle (selection-recombination-mutation-evaluation) is carried through several times; during these iterations the average fitness increases up to a certain threshold. In the end we reconsider the list of the individuals with the highest fitness value and choose the fittest among them as the final solution.

### B. Genetic algorithms in freezing problems

These general ideas of GAs are now adapted to describe freezing transitions: we start by representing a lattice—i.e., a possible candidate for a solid equilibrium structure—as an individual. We start with a simple lattice (i.e., with one basis particle per lattice site) and extend the formalism to more complex, nonsimple lattices (with more than one basis particle) later. Let $\{\mathbf{x}_i\}=\{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3\}$ be the primitive vectors of the Bravais lattice: then we use the following convention for the parametrization of the $\{\mathbf{x}_i\}$:

$$\mathbf{x}_1 = a(1,0,0),$$

$$\mathbf{x}_2 = a(x\cos\Phi, x\sin\Phi, 0),$$

$$\mathbf{x}_3 = a(xy\cos\Psi\cos\vartheta, xy\cos\Psi\cos\vartheta, xy\sin\vartheta). \tag{1}$$

In these equations, the parameter $a$ is uniquely determined by the density of the system; the remaining five parameters ($x$, $y$, $\Phi$, $\Psi$, and $\vartheta$), that characterize the Bravais lattice are limited by the following constraints:

$$0 < x \leqslant 1, \quad 0 < y \leqslant 1, \tag{2}$$

$$0 < \Phi \leqslant \pi/2, \quad 0 < \Psi \leqslant \pi, \quad 0 < \vartheta \leqslant \pi/2. \tag{3}$$

Note that representation (1) is *not* unique, since a given lattice can be represented by a different, but equivalent set of basis vectors $\{\overline{\mathbf{x}}_i\}$; we will reconsider this aspect later.

In nonsimple lattices with $b$ particles per basis, additional vectors $\mathbf{y}_i$, $i=1,\ldots,b$, have to be introduced, which fix the positions of the $b$ additional particles; in our calculations $b$ was limited due to practical reasons to 8 (see also discussion in Sec. IV). Without loss of generality we put $\mathbf{y}_1=\mathbf{0}$ and thus reduce the number of parameters to be optimized. For the $\mathbf{y}_i$, $i=2,\ldots,b$, we have used the following representation: $\mathbf{y}_i=\Sigma_{j=1,2,3}c_{ij}\mathbf{x}_j$, with $0\leqslant c_{ij}<1$. Again, note that different sets of $c_{ij}$ can characterize different but equivalent representations of a nonsimple lattice; also these ambiguities have been excluded by suitable algorithms—see below. Other parameters that characterize the model can be added if required (see, for instance, the examples discussed in Sec. III B).

We now "translate" a Bravais lattice, represented by $\{x,y,\Phi,\Psi,\vartheta\}$, into an individual $\mathcal{I}$ where we have chosen the following convention:

$$\{x_1,x_2,x_3\} \equiv \{x,y,\Phi,\Psi,\vartheta\} \rightarrow b_x b_y b_\Phi b_\Psi b_\vartheta = \mathcal{I}_{\{x_i\}}, \tag{4}$$

the $b_x,\ldots,b_\vartheta$ represent five strings of genes of different lengths: in our approach we have found 12 genes to the sufficient to build up $b_x$ and $b_y$ and 6 genes for $b_\Phi$, $b_\Psi$, and $b_\vartheta$; see also discussion below.

For the genes we have chosen the binary alphabet, therefore the $b_i$ are sequences of 0's and 1's. With this encoding the $b_i$'s can be interpreted as binary representations of a decimal number, which we denote by $\tilde{b}_i$. We further introduce $b_{i,\max}$ to denote the maximum value of the $b_i$ (being a sequence of 6 or 12 1's), and its decimal representation, $\tilde{b}_{i,\max}$ (being in our case either 4095 or 63).

The value of $x$ is obtained via

$$x = \frac{\tilde{b}_x + 1}{\tilde{b}_{x,\max} + 1}. \tag{5}$$

This relation between $x$ and $b_x$ has the attractive feature that it guarantees fulfillment of restriction (2) on $x$ and that unphysical (i.e., negative) values of $x$ are automatically excluded. Due to the limited length of $b_x$, results for $x$ have a relative accuracy of $2^{-12}\sim 2\times 10^{-4}$, a value that we have found to be sufficient in our studies. For $y$ we use exactly the same translation rule. For the angles $\Phi$, $\Psi$, and $\vartheta$ we use a slightly different parametrization, namely,

$$\Phi = \frac{\pi}{2}\frac{\tilde{b}_\Phi + 1}{\tilde{b}_{\Phi,\max} + 1} \tag{6}$$

and similar relations for $\Psi$ and $\vartheta$; for obvious reasons these expressions take into account restrictions (3) on the angles; the relative accuracy for these parameters is thus $2\times 10^{-6}\sim 0.016$.

In nonsimple lattices, the additional coordinates of the basis particles (i.e., the quantities $c_{ij}$) have to be taken into account: for each $c_{ij}$ we use a translation rule similar to Eq. (5); thus an individual $\mathcal{I}_{\{\mathbf{x}_i,\mathbf{b}_j\}}$, representing now a nonsimple lattice, has the length of $[18+12(3b-1)]$ genes. If we restrict ourselves to up to eight basis particles, the maximum length of an individual can reach in our case a value of 294; this means that the possible search space is extremely high dimensional. As will be discussed below (Sec. IV), the surface that describes the fitness of the individuals in search space becomes extremely rugged (i.e., characterized by many local minima) in high-dimensional search spaces, causing thus considerable problems in the search of the absolute mini-

mum. Therefore, the limitation to eight basis particles represents rather a numerical than a practical limit.

We now come back to the uniqueness of the representation of a lattice by an individual $\{\mathbf{x}_i,\mathbf{b}_j\}\leftrightarrow\mathcal{I}_{\{\mathbf{x}_i,\mathbf{b}_j\}}$. While in other problems in condensed matter physics it might be irrelevant if a lattice is represented by a set of primitive vectors $\{\mathbf{x}_i\}$ or by an equivalent set $\{\overline{\mathbf{x}_i}\}$, this is certainly not the case in an application of a GA. We explain this in a simple example: suppose that the face-centered-cubic (fcc) structure is the equilibrium structure for a given state point. If this fcc lattice were not represented by a uniquely fixed set of primitive vectors $\{\mathbf{x}_i\}$ then the GA might propose two (or perhaps even more) sets of equivalent primitive vectors that represent the very same fcc lattice as optimal solutions and the algorithm would never converge. Therefore we had to introduce a convention that fixes the set of the $\{\mathbf{x}_i\}$ and, possibly, of the set of the $\{\mathbf{b}_i\}$ in a unique way.

Again we start to present this convention with a Bravais lattice. First, we choose the primitive vectors $\{\mathbf{x}_i\}$ in such a way that the surface of the unit cell, $\Sigma$, be minimal. The latter is given by the equation:

$$\Sigma = |\mathbf{x}_1 \times \mathbf{x}_2| + |\mathbf{x}_1 \times \mathbf{x}_3| + |\mathbf{x}_2 \times \mathbf{x}_3|. \tag{7}$$

This can be achieved by a simple iterative algorithm where the vectors $\{\mathbf{x}_i\}$ are replaced by an equivalent set of $\{\mathbf{x}_i'\}$ until $\Sigma$ reaches its minimum. Starting with the three vectors $\{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3\}$ one checks the surface of the cell spanned by the 12 combinations,

$$\{\mathbf{x}_1 \pm \mathbf{x}_2,\mathbf{x}_2,\mathbf{x}_3\},\quad \{\mathbf{x}_1 \pm \mathbf{x}_3,\mathbf{x}_2,\mathbf{x}_3\},$$

$$\{\mathbf{x}_1,\mathbf{x}_2 \pm \mathbf{x}_1,\mathbf{x}_3\},\quad \{\mathbf{x}_1,\mathbf{x}_2 \pm \mathbf{x}_3,\mathbf{x}_3\},$$

$$\{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3 \pm \mathbf{x}_1\},\quad \{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3 \pm \mathbf{x}_2\},$$

and takes the combination with the lowest value as new primitive vectors if the surface is smaller than the one of the starting vectors. If the surface increases by combining the vectors the algorithm is ended.

The set of basis vectors $\{\mathbf{b}_i\}$ is also not unique, as the $(b-1)$ sets $\{\mathbf{b}_i-\mathbf{b}_j\}$, $j=2,\ldots,b$ describe the same lattice and the basis vectors can be ordered in any way without changing the lattice. We chose the following convention to fix the basis vectors: all basis vectors are represented by their coordinates with respect to the primitive vectors. The coordinates can be shifted by integer values until they lie in the interval $[0,1)$. The basis vector $\mathbf{b}_j$ is chosen so that the sum of the coordinates of all basis vectors is a minimum. The vectors are sorted first by their $\mathbf{x}_1$ coordinate, then by the $\mathbf{x}_2$ coordinate, and then by the $\mathbf{x}_3$ coordinate.

Second, taking advantage of our freedom with respect to rotation and/or inversion of the $\{\mathbf{x}_i\}$, the vectors are rotated until the longest of them points in the positive $x$ axis. The other two vectors are then rotated around the $x$ axis so that the second longest vector lies in the $(x,y)$ plane. If necessary, the shortest vector and/or the $z$ axis are inverted so that the $x$ and the $z$ component of the shortest vector is positive. One then arrives at the representation (1). Finally, based on the results for the $\{\mathbf{x}_i\}$, the lattices were identified by analyzing their symmetry properties.

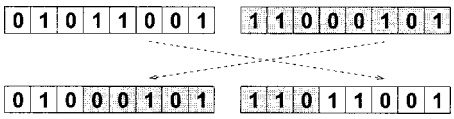


FIG. 1. Schematic representation of the one-point crossover that generates from two parent individuals, $\mathcal{I}'_{[i]}$ and $\mathcal{I}''_{[i]}$, of generation $\mathcal{G}_{[i]}$ (top) and two new individuals, $\mathcal{I}'_{[i+1]}$ and $\mathcal{I}''_{[i+1]}$, of the subsequent generation $\mathcal{G}_{[i+1]}$ (bottom).

At this occasion we point out that the GA is self-consistent with respect to the representation of the structure. We explain this for the example of a fcc structure: if for a given state point the fluid freezes into a fcc structure it is irrelevant if this lattice is represented by a simple lattice or by a non-simple lattice: the GA proposes via the different routes within high accuracy the *same* fcc structure.

We have implemented our GA as follows: we start with the generation of index 0, $\mathcal{G}_{[0]}=\{\mathcal{I}_{[0]}\}$, where the individuals $\{\mathcal{I}_{[0]}\}$ are chosen at random. Typical values for the number of individuals, comprising a generation $N_{\mathcal{I}}$ used in our GAs are 1000 (see discussion in Sec. IV). As already outlined in Sec. II A, a positive fitness value is assigned to each individual $\mathcal{I}$ via the fitness function $f(\mathcal{I})$: a higher fitness value characterizes a better solution. For our problem, we specify the fitness function as follows: since the crystal structure with the lowest free energy $F$ is considered to be the stable one, the fitness function in our approaches has to be energy related. In our investigations the standard choice for $f(\mathcal{I})$ reads

$$f(\mathcal{I}) = \exp\left\{1 - \frac{F(\mathcal{I})}{F(\mathcal{I}^{\mathrm{fcc}})}\right\}, \tag{8}$$

where $F(\mathcal{I}^{\mathrm{fcc}})$ is the free energy of a fcc structure. The appearance of the ratio $F(\mathcal{I})/F(\mathcal{I}^{\mathrm{fcc}})$ in the exponent guarantees that the latter always remains of order unity, since both quantities involved are extensive. Alternative functional forms will be discussed in Sec. IV.

Subsequently, we create in an iterative process from the individuals of generation $\mathcal{G}_{[i]}$ the individuals of the subsequent generation $\mathcal{G}_{[i+1]}$: we start by selecting two individuals, $\mathcal{I}'_{[i]}$ and $\mathcal{I}''_{[i]}$, according to their fitness values as parents. The probability that an individual $\mathcal{I}_{[i]}$ is chosen as parent is given by

$$p(\mathcal{I}_{[i]}) = \frac{f(\mathcal{I}_{[i]})}{\sum_{\mathcal{J}\in\mathcal{G}_{[i]}} f(\mathcal{J})}. \tag{9}$$

The two chosen parents create, in the second step, individuals $\mathcal{I}'_{[i+1]}$ and $\mathcal{I}''_{[i+1]}$ of the subsequent generation $\mathcal{G}_{[i+1]}$, via a one-point crossover, as visualized in Fig. 1. To this end, we cut the parents' gene sequences at a randomly chosen point and cross combine them. A uniform random integer number $r\in[1,l-1]$ determines the crossover point, where $l$ denotes the length of the individual. We point out that this recombination step is completely "blind" to the cell geometry (in terms of the parameters $x$, $y$, $\Phi$, $\Psi$, and $\vartheta$) and the location of the (possible) basis atoms. It therefore *does not* correspond to geometric moves of particles which thus clearly

distinguishes our GA approach from applications of evolutionary algorithms to describe cluster formation.[18]

Further, we perform mutations with a probability $p_{\mathrm{mutate}}$ (which has typically a value of 0.01) on the individuals $\mathcal{I}_{[i+1]}$: this means that we flip in an arbitrarily chosen individual in an arbitrarily position the value of a gene from zero to one or vice versa. Mutations avoid inbreeding and represent at the same time reintroduction of new or lost genetic information; they will be discussed thoroughly in Sec. IV. Finally, we evaluate the genes that have been constructed in this way via the fitness function, $f(\mathcal{I}_{[i+1]})$, and denote the one with the highest fitness value in generation $\mathcal{G}_{[i+1]}$ by $\mathcal{I}^*_{[i+1]}$. We then iterate this cycle, producing thereby a typical number of $N_{\mathcal{G}} \sim 100$ generations. The $\mathcal{I}^*_{[i]}$ are recorded and the final solution of the GA, $\mathcal{I}^*$, is considered to be the individual among the $\mathcal{I}^*_{[i]}$ with the highest fitness value. The final solution of the problem is found by a final hill-climbing optimization, starting from $\mathcal{I}^*$; we need this final step in order to compensate for the limited accuracy of the parameters $\{x, y, \Phi, \Psi, \vartheta\}$ which is a consequence of the limited number of genes in the respective binary representations, Eqs. (5) and (6).

## III. EXAMPLES

### A. Star polymers

Star polymers are typical representatives of soft matter systems; they are complex aggregates, built up by a well-defined number $f$ of flexible polymer chains that are grafted to a central particle. In Ref. 23 an effective potential $\Phi(r)$ has been proposed, which was obtained by averaging over the many degrees of freedom of the fluctuating chain monomers via suitable coarse graining methods; it reads as

$$\beta\Phi(r) = \frac{5}{18} f^{3/2} \begin{cases} -\ln(r/\sigma) + (1+\sqrt{f}/2)^{-1}, & r \leqslant \sigma \\ (\sigma/r)(1+\sqrt{f}/2)^{-1} \exp[-\sqrt{f}(r-\sigma)/2\sigma], & \sigma < r. \end{cases} \tag{10}$$

Two parameters characterize the system: the functionality $f$ and the corona diameter $\sigma$, which measures the spatial extension of the monomer density around the central particle. $\beta = (k_{\mathrm{B}}T)^{-1}$, $k_{\mathrm{B}}$ is Boltzmann's constant, $T$ is the temperature, and $\rho$ is the number density. Because of the purely entropic nature of the model that leads to above potential,[23] $\Phi(r)$ scales with temperature which thus becomes an irrelevant parameter.

In work subsequent to Ref. 23, the phase behavior of star polymer particles was determined, using the effective potential of Eq. (10) as a starting point.[3,24] We focus in the following on the solid phases only. To be consistent with Ref. 3, we restrict ourselves, as we calculate the energies of each of the competing solid structures, to lattice sums $L$: this means that the particles are considered to be fixed at the lattice positions and that no thermal fluctuations are taken into account. Phase diagrams based on lattice sums are commonly referred to in literature as the "zero-temperature phase diagram" since particles are considered to be immobile in their lattice positions; even though temperature is an irrelevant parameter in the present model we shall use this term in this contribution as well. Note that at $T=0$, $F=U=L$.

In the conventional approach to determine crystal equilibrium structures one starts by a preselection of possible (or plausible) candidate structures. Since the phase diagram of a soft system has been in the late 1990s a yet undiscovered field, the authors of Refs. 3 and 24 had to rely in their search for possible structures on their intuition or on plausible arguments and particular care was in order. Apart from the obvious candidate, fcc and body centered cubic (bcc), also body-centered orthogonal (bco) and diamond lattices were found to be stable structures. This represents already a remarkable and completely unexpected result since neither bco (having an anisotropic rectangular cell) nor diamond were expected as equilibrium structures for a system with radially symmetric pair potential: there was rather a widespread belief that such structures can only be induced by nonspherical potentials. Additional and very careful considerations, which, in turn, were based on a rather general study on the driving mechanism of the open structures in soft systems,[20,21] led to the conclusion that the A15 structure might also represent a possible candidate structure for star polymers; it was indeed encountered for rather high densities. In Fig. 2 we display the conventional unit cell of this rather rare structure. The complete zero-temperature phase diagram, being depicted in Fig. 6 of Ref. 3, is based on these five candidates and has quite an exotic topology: it is characterized by a host of many equilibrium structures, it has an unexpectedly complex topology, including both first- and second-order phase transitions as well as critical end points.

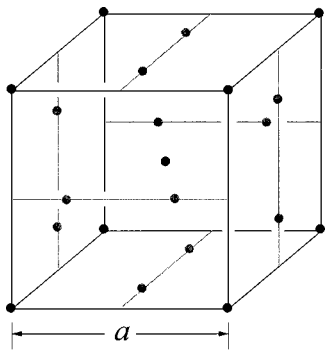


FIG. 2. The conventional unit cell of the A15 structure. Note that this cell can be described by a simple cubic lattice plus eight basis particles. The length of the conventional cell edge is denoted by $a$.

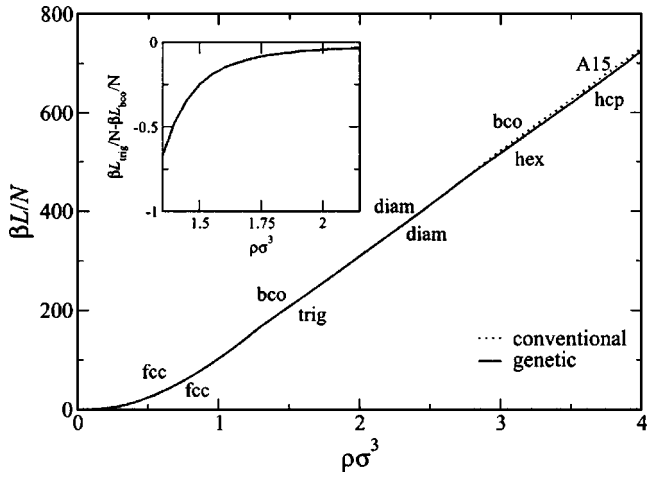


FIG. 3. Dimensionless lattice sum per particle $\beta L/N$ for star polymers with functionality $f=64$ as a function of the density $\rho\sigma^3$, evaluated for the equilibrium structures found in the conventional approach (Ref. 3) (broken line) and for the equilibrium structures predicted by the GA (present contribution; full line). The inset shows the difference in $\beta L/N$ between the hexagonal and the trigonal structure for $1.3 \lesssim \rho\sigma^3 \lesssim 2.2$.

We have reconsidered the very same problem, using now our GA to find the equilibrium structures of the solid phases. We shall demonstrate that—in contrast to the conventional approach outlined above—the GA is able to predict the solid equilibrium structures in a parameter-free, unbiased, and unrestricted search. Since we use lattice sums $L$ to calculate the zero-temperature phase diagram, the free energy $F$ in expression (8) for the fitness function $f(\mathcal{I})$ is replaced by $L$, which is given by

$$L = \frac{1}{2}\sum_{\mathbf{R}_i \neq 0} \Phi(\mathbf{R}_i), \tag{11}$$

where $\{\mathbf{R}_i\}$ are the vectors denoting the particle positions at the given lattice. At each state (characterized by $\rho$ and $f$) the GA predicts, the respective solid equilibrium structure.

In Fig. 3 we display the lattice sum $L$ as a function of $\rho$ for star polymers with functionality $f=64$; we compare the structures obtained in the conventional approach[3] with the equilibrium structures predicted via the GA. Already at intermediate densities a trigonal lattice, which had not been considered in the conventional approach, is obviously more stable than bco. However, as can be seen from the inset of Fig. 3, the difference in the free energies is of the order of $\sim 0.5\%$: this convincingly demonstrates the difficulties in the conventional approach and the attractive features of the GA. As we increase the density, both approaches predict a diamond structure. For $\rho\sigma^3 \gtrsim 3$, additional and this time remarkable discrepancies between the two approaches are observed: the GA predicts hexagonal and hcp as equilibrium structures, i.e., two crystals that had not been considered in the conventional approach at all; for these densities bco and A15 were predicted in Ref. 3. We summarize that for $\rho\sigma^3 \lesssim 4$ the GA predicts three equilibrium structures that had not been considered in the previous approach, i.e., trigonal, hcp, and hexagonal, all of them giving lower free energies than the ones originally proposed in Ref. 3.

For $4.1 \lesssim \rho\sigma^3 \lesssim 4.4$, the GA predicts—in agreement with the results presented in Ref. 3—the A15 structure (see Fig. 2), a rather exotic lattice that has been encountered experimentally in soft matter.[25] We point out that the fact that the GA is able to predict the A15 as a possible equilibrium structure in this complex and highly dimensional search space (note that an individual is now built up by 252 genes and the search space comprises $2^{252} \sim 7 \times 10^{75}$ candidate structures) convincingly demonstrates the power of this approach. We close with the remark that the GA has, as expected, convergence problems close to the phase boundaries, where the two competing structures have values of the free energy which are very close.

### B. Charged microgels

In the next example we consider charged microgels: in such a system we find charged, spherical macroions and point counterions dispersed in an electrolyte solvent; the latter is treated as a continuum and is characterized by a dielectric constant $\epsilon$. Denton[26] has proposed a simple model for this system and has derived, by averaging over the degrees of freedom of the constituent particles, explicit expressions for the effective potential $\Phi(r)$ acting between the centers of the microgel particles and for the thermodynamic properties. Written in a compact form, $\Phi(r)$ is given by

$$\beta\Phi(r) = \frac{Z^2\lambda_B}{\sigma}\begin{cases} \left\{\frac{24}{\kappa^2\sigma^2} - \frac{72}{\kappa^3\sigma^3}\left[e^{-\kappa\sigma}\left(1+\frac{2}{\kappa\alpha}\right)^2 + \left(1-\frac{4}{\kappa^2\sigma^2}\right)\right]\right\}, & r=0 \\ \left\{\frac{24}{\kappa^2\sigma^2} + \frac{r}{\sigma}\left(\frac{144}{\kappa^4\sigma^4} - \frac{36}{\kappa^2\sigma^2}\right) + \frac{r^3}{\sigma^3}\frac{12}{\kappa^2\sigma^2} - \frac{72}{\kappa^4\sigma^4}\frac{\sigma}{r}\left[e^{-\kappa\sigma}\left(1+\frac{2}{\kappa\sigma}\right)^2 \sinh(\kappa r) + \left(1-\frac{4}{\kappa^2\sigma^2}\right)(1-e^{-\kappa r})\right]\right\}, & 0<r\leq\sigma \\ \frac{144}{\kappa^4\sigma^4}\frac{\sigma}{r}\left\{\cosh(\kappa\sigma/2) - \frac{2}{\kappa\sigma}\sinh(\kappa\alpha/2)\right\}^2 e^{-\kappa r}, & \sigma<r. \end{cases} \tag{12}$$

In Eq. (12) above, $Z$ is the effective macroion valence, $\lambda_B$ is the Bjerrum length, and $\kappa=\sqrt{4\pi n_c z^2 \lambda_B}$ is the inverse Debye screening length, where $n_c$ stands for the counterion number density and $z$ for their valency; $\sigma$ is the diameter of the macroions.

The phase diagram of charged microgels has already been presented in Refs. 27 and 28. Here we complement these data with a few remarks on technical modifications of the GA for this particular system. In contrast to the star polymers, we have calculated the finite-temperature phase diagram via the Einstein model[8,29] (for reasons given below); this approximate concept provides an estimate for the upper bound of the free energy of a system. Its basic relation is the Gibbs–Bogoliubov inequality that relates the free energy of the system, $F$, and the free energy of a suitably chosen reference system, $F_0$, via the inequality

$$F \leqslant F_0 + \langle \mathcal{V} - \mathcal{V}_0 \rangle_0, \tag{13}$$

$\mathcal{V}$ and $\mathcal{V}_0$ are, respectively, the potential energies of the system and of the reference system. $\langle \cdots \rangle_0$ represents an ensemble average taken in the reference system for which we chose an Einstein solid: here the potential energy is given by

$$\mathcal{V}_0 = \sum_{\{\mathbf{R}_i\}} \frac{\alpha}{2} (\mathbf{r}_i - \mathbf{R}_i)^2. \tag{14}$$

$\alpha$ is a spring constant and the $\{\mathbf{R}_i\}$ represent the positions of the lattice. For $F_0$ one can derive the following result:

$$\beta \frac{F_0}{N} = \frac{3}{2} \ln\left(\beta \frac{\alpha \sigma^2}{2\pi}\right) + 3 \ln\left(\frac{\Lambda}{\sigma}\right), \tag{15}$$

where $\Lambda$ is the de Broglie wave length. Finally, $\mathcal{V}$ is given by

$$\mathcal{V} = \sum_{\{\mathbf{R}_i, \mathbf{R}_j\}} \Phi(\mathbf{R}_i - \mathbf{R}_j). \tag{16}$$

For charged microgels the evaluation of $\langle \mathcal{V} - \mathcal{V}_0 \rangle_0$ can be carried out *analytically*; closed expressions are presented in Ref. 30. This nice feature was not accessible for the potential of star polymers due to the logarithmic term in Eq. (10). We calculated both the zero-temperature and the finite-temperature phase diagrams for the solid phases obtained by the GA. Both results are depicted in Fig. 4.

Two remarks are of order: first, we note that if we use lattice sums instead of the Einstein model to calculate the free energy, then we obtain a qualitatively similar phase diagram with exactly the same candidate structures; the precise locations of the phase boundaries are of course different in the two approaches. We further point out that the Einstein model becomes equivalent to a density-functional theory (DFT) approach, where a mean-field format for the excess free-energy functional $\mathcal{F}_{ex}[\varrho]$ and a Gaussian-shaped one particle density $\varrho(\mathbf{r})$ is used:[31]

$$\mathcal{F}_{ex}[\varrho] = \frac{1}{2} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \varrho(\mathbf{r}_1) \varrho(\mathbf{r}_2) \Phi(|\mathbf{r}_1 - \mathbf{r}_2|), \tag{17}$$

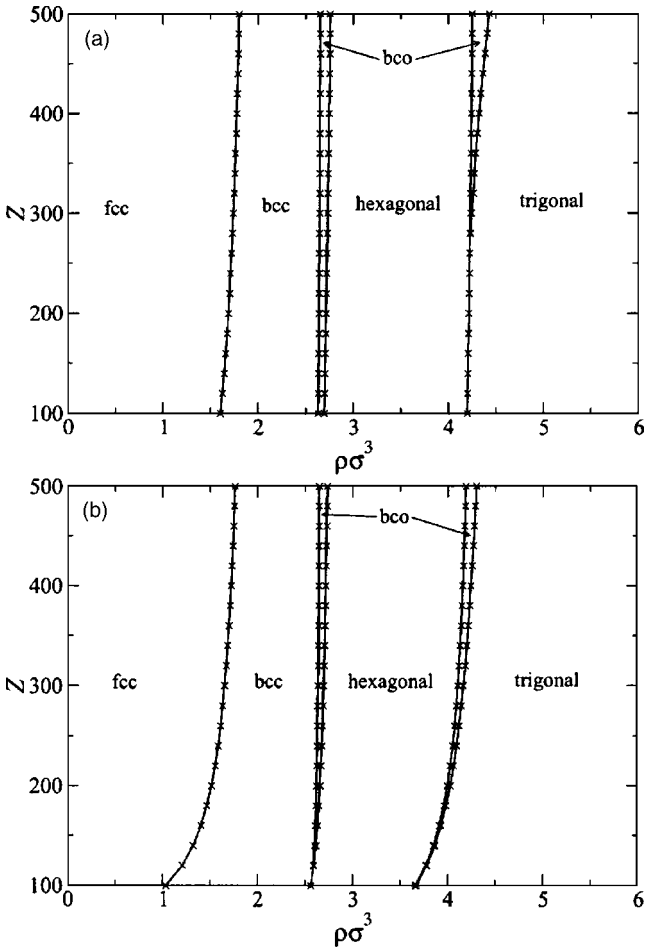


FIG. 4. (a) Zero-temperature solid phase diagram for ionic microgels. (b) Finite temperature solid phase diagram for ionic microgels. In both diagrams the microgel diameter is fixed at $\sigma=100$ nm and the phase boundaries are the intersection points of the free-energy curves, i.e., the small density gaps between coexisting phases are not shown.

$$\varrho(\mathbf{r}) = \left(\frac{\tilde{\alpha}}{\pi}\right)^{3/2} \sum_{\{\mathbf{R}_i\}} \exp[-\tilde{\alpha}(\mathbf{r} - \mathbf{R}_i)^2], \tag{18}$$

where the $\{\mathbf{R}_i\}$ are the lattice positions. Once can show[31] that then the expression for the free energy (as a function of $\tilde{\alpha}$) in the DFT approach is exactly the right-hand side of inequality Eq. (13) in the Einstein model. Minimization of the free energy with respect to $\tilde{\alpha}$ in the DFT approach leads to the equilibrium one particle density; minimization of the right-hand side of Eq. (13) with respect to $\alpha$ provides an estimate for the smallest upper bound for the free energy of the system. Thus the two approaches are equivalent.

This minimization of the right-hand side of Eq. (13) with respect to the set of parameters of the reference system (in our case, the parameter $\alpha$) in the Einstein model can be integrated in the GA formalism. To this end we have extended the representation of an individual—see Eq. (4) for the simple lattice—by an additional sequence of genes, $b_\alpha$, which is a string of ten genes, of them assuming again values 0 or 1. Introducing a reasonable lower limit for $\alpha$, i.e., $30 \leqslant \alpha$, prevents that the delocalization of the particles becomes unphysically large. Again, this restriction on $\alpha$ can be suitably included in the formalism: we relate $\alpha$ and $\tilde{b}_\alpha$, the value

of the decimal representation of $b_\alpha$, via $\alpha=30+\tilde{b}_\alpha n_\alpha$; we have chosen $n_\alpha=5$, and thus $\alpha\in[30,5145]$. Although formally $\alpha$ can assume only values on a grid of integer number with spacing of $n_\alpha$, this does not represent a severe restriction: the final value for $\alpha$ proposed by the GA is refined in the subsequent hill-climbing search.

The zero-temperature phase diagram obtained in this way is depicted in Fig. 6 of Ref. 28 and contains for $\rho\sigma^3<6$ the following equilibrium structures (in the sequence of increasing density): fcc, bcc, hexagonal, bco (only for a number of elementary charges $Z\gtrsim 450$), trigonal, and, again, hexagonal. In our earlier investigations of the charged microgels,[27,28] we have calculated the complete phase diagram for $T>0$ (including both solid *and* liquid phases). For this task the GA has fulfilled an extremely useful task, since the structures obtained for the zero-temperature phase diagram via the GA have now been used as possible candidates for the the solid structures of the complete phase diagram.

## IV. REFINEMENTS OF THE GA AND THE ROLE OF NUMERICAL PARAMETERS

In this section we propose refinements of the basic version of the GA as presented and applied in previous sections; we also discuss the influence of the numerical parameters on the convergence and efficiency of the algorithm. The examples we use for demonstration are typical and representative cases that we have encountered; since in our GA version the individuals of the first generation are created via random numbers and the subsequent crossover-, mutation-, and selection-processes are also controlled by random numbers, the curves we present in the following are reproducible only on a qualitative but not on a quantitative level. For convenience we introduce the notation $L^*_{[i]}=L(\mathcal{I}^*_{[i]})$, i.e., the lattice sum of the fittest (best) individual of generation with index $i$.

### A. Refinements of the GA technique

In the preceding sections we have introduced the basic version of the GA. Due to several reasons, refinements of this technique are of order. First, they aim at a higher convergence speed: the evaluation of the fitness function (or, equivalently, the free energy/the lattice sum) represents the numerical bottleneck of the algorithm. Therefore, if we could obtain the final solution of the GA with a considerably smaller number of individuals and/or less generations, using refined concepts to calculate the free energy, more complex applications might come within reach. Another important issue is the dimensionality of search space: we have realized that from a certain number of basis vectors onwards (typically eight vectors or more) the surface in search space becomes very rugged: it is then characterized by a large number of local minima and consequently the search of the *global* minimum becomes very difficult. Therefore, a reasonably reduction of the dimensionality of the search space (e.g, via a suitable bias) might be very helpful. Both aspects will be addressed in the following.

Convergence can indeed be sped up by introducing a generation-dependent fitness function, for example, via the following functional form:

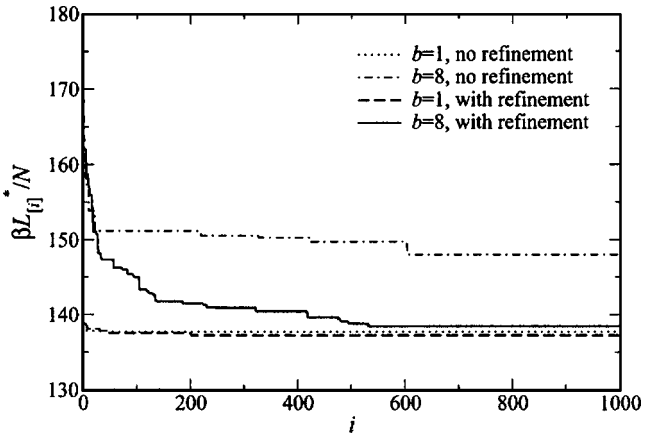


FIG. 5. Dimensionless lattice sum per particle of the fittest individual, $\beta L^*_{[i]}/N$, of the generation with index $i$ as a function of $i$, obtained via the fitness function $f(\mathcal{I})$, Eq. (8) (no refinement), and the modified fitness function $\tilde{f}(\mathcal{I})$, Eq. (19) (refinement), for a microgel system that freezes into a hexagonal structure ($Z=500, \sigma=100$ nm, $\rho\sigma^3=5$); monoatomic Bravais lattice ($b=1$) vs structure with eight basis particles ($b=8$).

$$\tilde{f}_{[i]}(\mathcal{I})=\exp\left\{1-\left[\frac{F(\mathcal{I})}{F(\mathcal{I}^{\mathrm{fcc}})}\right]^{\epsilon(i).}\right\}, \tag{19}$$

where the function in the exponent, $\epsilon(i)$, now depends on the generation index $i$; in the present contribution we have used $\epsilon(i)=1+i/10$. Thus the choice of the fittest individual in a generation becomes more selective as $i$ increases. To demonstrate the influence of the function $\epsilon(i)$ on the convergence of the GA we have chosen a state point, where a microgel fluid freezes into a hexagonal structure ($Z=500, \sigma=100$ nm, $\rho\sigma^3=5$). A total number of four calculations were carried out with the GA: A lattice with one basis vector ($b=1$) both with the fitness function (8) (no refinement) and the generation-dependent fitness function (19) (refinement) was optimized and two analogous calculations for a lattice with eight basis particles ($b=8$). In Fig. 5 we display the lattice sum of the fittest individual of generation $i$, $L^*_{[i]}$, as a function of the generation index $i$.

The fact that the GA leads after 1000 generations to results that differ by $\sim 10\%$ might seem puzzling at first sight, since all routes should converge to the same structure; in fact, the final hill-climbing search *does* lead to the same final result ($\beta L_{\mathrm{final}}/N\sim 137.13$). In the case where eight basis particles and the refined fitness function were used, the primitive vectors form a face-centered monoclinic lattice. From the four different curves presented in Fig. 5 we learn the following: after 1000 generations the fitness function $\tilde{f}(\mathcal{I})$ provides for both $b$ values lower results for $L^*_{[i]}$ than the function $f(\mathcal{I})$ and proposes thus a better starting point for the hill-climbing search. While these differences are rather negligible for the low-dimensional search space ($b=1$), they are considerable for $b=8$; this is a very encouraging message since the improved fitness function brings along a significant improvement, in particular, for the case where—due to the large number of basis particles—the search space is extremely high dimensional and where the search for a minimum is very delicate—as mentioned above. In addition, this nice advantage is provided for free since the evaluation of

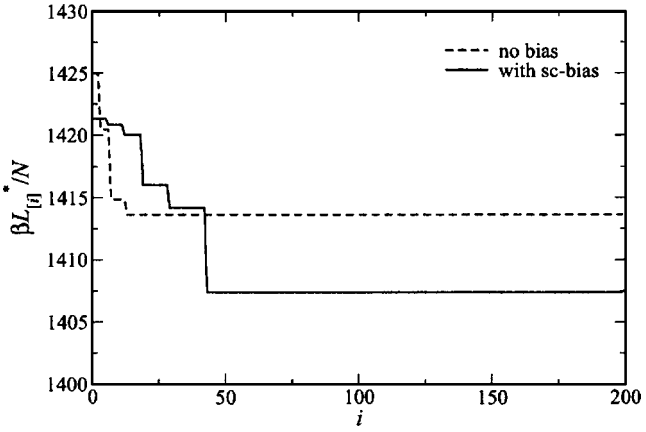


FIG. 6. Dimensionless lattice sum per particle of the fittest individual, $\beta L^*_{[i]}/N$, of generation with index $i$ as a function of $i$: conventional GA (broken line) lattice (sc) biased version (full line).

the fitness value via Eq. (19) is not more time consuming than via Eq. (8); thus no additional computational effort has to be made to obtain this improved convergence.

The other modification of the algorithm we propose is a lattice-biased version of the GA. It is designed to bring along improvements for those cases when the underlying Bravais lattice of the equilibrium structure is obvious, while the number and the positions of the basis particles is yet unknown. Then one can fix the vectors $\{\mathbf{x}_i\}$ during the search procedure and vary only the number and the parameters that fix the positions of the basis. This leads to a reduction in the number of the parameters to be optimized and thus to a decrease in the computational effort. We demonstrate the consequences of the lattice-biased version of the GA in Fig. 6 with the following example: we have chosen a state point where the star polymer fluid freezes into the A15 structure; the conventional unit cell has cubic symmetry and a basis of eight particles (cf. Fig. 2 and the discussion in Sec. III A). In Fig. 6 we display the dimensionless lattice sum per particle of the fittest individual of generation with index $i$, $\beta L^*_{[i]}/N$, calculated via the conventional GA and via the lattice-biased version of the algorithm; in the latter case, a simple cubic (sc) lattice was assumed with fixed lattice vectors and only the positions for the eight particles of the basis were optimized. Even though the conventional algorithm converges faster, the lattice-biased version leads to a lower result for the lattice sum and thus to a better starting value for the subsequent hill-climbing search.

### B. The role of the numerical algorithm parameters

In this section we turn our attention to the role played by the numerical parameters of the algorithm and present some details regarding the choice of their precise values. In the course of the present investigation but also in dealing with other applications, such as clustering transitions and quasi-two-dimensional systems we have come to the conclusion that the optimal choice of the numerical parameters ($N_{\mathcal{I}}$, $N_{\mathcal{G}}$, length of binary representations, etc.) are strongly problem specific and therefore no general guidelines can be provided. For the present contribution, $N_{\mathcal{I}}=1000$ and $N_{\mathcal{G}}=100$ have

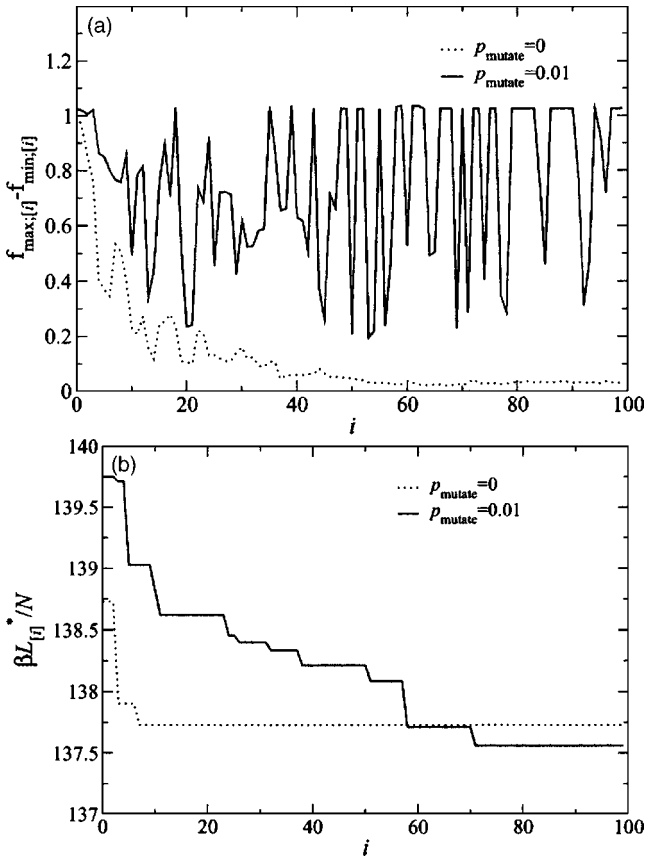


FIG. 7. (a) The quantity $\Delta f_{[i]}=f_{\max;[i]}-f_{\min;[i]}$ as a function of the generation index $i$, where $f_{\max;[i]}$ ($f_{\min;[i]}$) is the highest (lowest) fitness value encountered in generation with index $i$. Full line, ($p_{\text{mutate}}=0.01$); dotted line, ($p_{\text{mutate}}=0$). (b) The dimensionless lattice sum per particle of the fittest individual, $\beta L^*_{[i]}/N$, of generation with index $i$ as a function of $i$; lines as in (a).

turned out to be largely sufficient, for most of the state points studied convergence was already obtained after a considerably smaller number of generations.

We therefore strongly recommend to test thoroughly these numerical parameters in practical applications. Further, computational effort can be economized by using and/or adapting the refined concepts introduced in the preceding section. In particular, modifications of the fitness function following the lines of expression (19) have turned out to be extremely useful in high dimensional search spaces. We point out once more that the (problem-specific) evaluation of the fitness function represents the bottleneck of the GA: most of the computational effort can be economized here.

Let us now consider the mutation step in the GA and demonstrate in the following that it represents a very important part of the GA. In Fig. 7(a) we display $\beta L^*_{[i]}/N$ as a function of the generation index $i$: once it has been calculated with a mutation probability $p_{\text{mutate}}=0.01$, then we have suppressed the mutation completely in the GA (i.e., $p_{\text{mutate}}=0$). The curves clearly demonstrate that mutation is an indispensable ingredient to improve the quality of the fitness of the individual as the generation develops: for $p_{\text{mutate}}=0$, the GA leads after a few generations ($i\cong 6$) to an individual with a relatively small value for the lattice sum and from there onwards the quality of the individuals does not improve any

more. On the other hand, for $p_{\mathrm{mutate}}=0.01$ there is a steady improvement in the lattice sum of the fittest individual as $i$ increases over nearly 70 generations. The influence of mutation on the genetic diversity can be seen in Fig. 7(b): here we display $\Delta f_{[i]}$, i.e., the difference in the highest ($f_{\max;[i]}$) and the lowest ($f_{\min;[i]}$) fitness values encountered in the generation with index $i$. While for $p_{\mathrm{mutate}}=0$, $\Delta f_{[i]}$ soon reaches values close to zero ($\Delta f_{[i]} \lesssim 0.1$ for $i \gtrsim 30$), the genetic diversity in the individuals over the generations for $p_{\mathrm{mutate}}>0$ is clearly visible by the relatively large $\Delta f_{[i]}$ values and their strong variation as a function of $i$. We point out that these results nicely reproduce those findings that were reported on the mating behavior of *Parus caeruleus*[22] mentioned in the introduction: for those birds the extra pair mating (which corresponds in our case to the mutation step) leads to a genetically diversified offspring, which has higher survival chances and an increased fitness.

## V. SUMMARY AND CONCLUDING REMARKS

We have introduced GAs as a new tool to a standard problem in condensed matter theory, which—despite considerable efforts in past years—is still a nontrivial one: the prediction of equilibrium structures in freezing transitions. Since such transitions have already been thoroughly studied over past decades in systems with harshly repulsive systems, we have rather focused on soft systems, where the potentials diverge weakly or even remain finite at the origin. The choice for this class of systems is also justified by the fact that here the stability of the candidate crystal structures has been much less investigated and is still poorly understood; in addition, first investigations in this highly actual and challenging field have indicated that the phase diagrams of soft systems offer a surprisingly rich wealth of new and unexpected equilibrium structures. We were thus convinced that these systems represent a very stringent and reliable test of the GA. In our implementation the individuals are binary representations of the primitive vectors of the lattice and the positions of possible, additional basis atoms. Based on an energy related fitness function, the typical GA cycle "selection-recombination-mutation-evaluation" has led after a reasonably small number of iterations to the final result of the GA. The limited accuracy in the binary representation of the lattice is compensated by a concluding refinement step, realized in a hill-climbing search. In our approach, the calculation of the fitness function, representing the numerical bottleneck of the algorithm, has been restricted to simple concepts to calculate the free energy, such as lattice sums, Einstein models, or simple density functionals. We have applied the GA in this contribution to star polymers and charged microgels, two typical soft matter systems. In these examples we have demonstrated the attractive features of this algorithm, i.e., its power, efficiency, flexibility, and simplicity. We have compared the performance of GAs to conventional algorithms, which, instead, have to rely in their prediction of possible candidate structures on the researcher's experience, intuition, or on plausible arguments; the GA, on the other hand, predicts these structures in a parameter-free, unbiased, and unrestricted search in the entire search space and is thus clearly superior to the conventional approach. The calculation of the fitness function, representing the bottleneck of the algorithm, has been restricted to simple concepts to calculate the free energy, such as lattice sums, Einstein models, or simple density-functional formats.

Further, we have presented conceptual and numerical refinements of GAs which aim at an increased convergence speed and on numerical stability, thus offering access to more complex approaches and to more refined schemes to calculate the free energy. We have shown that modifications of the fitness function and a lattice-biased version of the GA can lead to a considerable increase in the convergence speed; we have also discussed the numerical parameters and, in particular, the importance of the mutation step.

With the appropriate modifications, genetic algorithms can also be applied to a related but much more complex problem, namely, to the determination of ordered structures in confined systems, lying between two and three dimensions. Preliminary results in this direction reinforce our claim regarding the enormous predictive power of these tools. Since up to now the power and the attractive features of GAs have not been sufficiently recognized neither in physics nor in chemistry (while they have proven to be very attractive in many other fields), we hope to motivate with the present contribution condensed matter theoreticians to a more widespread use of this attractive, useful, and easy-to-handle tool.

## ACKNOWLEDGMENTS

The authors thank Professor Rainer Dirl (Wien) for helpful discussions. This work was supported by the Österreichischer Forschungsfonds under Project Nos. P14371-TPH, P15785-TPH, and P17823-N08, as well as by the Deutsche Forschungsgemeinschaft through the Sonderforschungsbereich SFB-TR6, "Physics of Colloidal Dispersions in External Fields," Subproject C3. Financial support by the Hochschuljubiläumsstiftung der Stadt Wien under Project No. H-1080/2002 is gratefully acknowledged.

[1] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic, New York, 1986).
[2] C. N. Likos, Phys. Rep. **348**, 267 (2001).
[3] C. N. Likos, N. Hoffmann, H. Löwen, and A. A. Louis, J. Phys.: Condens. Matter **14**, 7681 (2002).
[4] J. G. Kirkwood, J. Chem. Phys. **18**, 380 (1950).
[5] W. W. Wood, J. Chem. Phys. **20**, 1334 (1952).
[6] A. Bonissent, P. Pieranski, and P. Pieranski, Philos. Mag. A **50**, 57 (1984).
[7] K. W. Wojciechowski, Phys. Lett. A **122**, 377 (1987).
[8] C. F. Tejero, A. Daanoun, H. N. W. Lekkerkerker, and M. Baus, Phys. Rev. E **51**, 558 (1995).
[9] H. R. Glyde and G. H. Keech, Ann. Phys. **127**, 330 (1980).
[10] D. Oxtoby, in *Les Houches, Session LI, Liquids, Freezing and Glass Transition*, edited by J.-P. Hansen, D. Levesque, and J. Zinn-Justin (North-Holland, Amsterdam, 1991).
[11] Y. Singh Phys. Rep. **207**, 351 (1991).
[12] A. D. J. Haymet, in *Fundamentals of Inhomogeneous Liquids*, edited by D. Henderson (Marcel Dekker, New York, 1992).
[13] H. Löwen, Phys. Rep. **237**, 249 (1994).
[14] M. Schmidt, J. Phys.: Condens. Matter **15**, S101 (2003).
[15] B. Groh and S. Dietrich, Phys. Rev. E **63**, 021203 (2001).
[16] J. H. Holland, *Adaptation in Natural and Artificial Systems* (The University of Michigan Press, Ann Arbor, 1975).
[17] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[18] D. M. Daeven and K. M. Ho, Phys. Rev. Lett. **75**, 288 (1995); S. Darby,

T. V. Mortimer-Jones, R. L. Johnston, and C. Roberts, J. Chem. Phys. **116**, 1536 (2002).
[19]S. H. L. Klapp, D. J. Diestler, and M. Schoen, J. Phys.: Condens. Matter **16**, 7331 (2004).
[20]P. Ziherl and R. D. Kamien, Phys. Rev. Lett. **85**, 3528 (2000).
[21]P. Ziherl and R. D. Kamien, J. Phys. Chem. **105**, 10147 (2001).
[22]K. Foerster, K. Delhey, A. Johnsoen, J. T. Lifjeld, and B. Kempenaers, Nature (London) **425**, 714 (2003).
[23]C. N. Likos, H. Löwen, M. Watzlawek, B. Abbas, O. Jucknischke, J. Allgaier, and D. Richter, Phys. Rev. Lett. **80**, 4450 (1998).
[24]M. Watzlawek, C. N. Likos, and H. Löwen, Phys. Rev. Lett. **82**, 5289 (1999).
[25]V. S. K. Balagurusamy, G. Ungar, V. Percec, and G. Johansson, J. Am. Chem. Soc. **119**, 1539 (1997).
[26]A. R. Denton, Phys. Rev. E **67**, 011804 (2003); **68**, 049904(E) (2003).
[27]D. Gottwald, C. N. Likos, G. Kahl, and H. Löwen, Phys. Rev. Lett. **92**, 068301 (2004).
[28]D. Gottwald, C. N. Likos, G. Kahl, and H. Löwen, J. Chem. Phys. **122**, 074903 (2005).
[29]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt Saunders, Philadelphia, 1976).
[30]D. Gottwald, Ph.D. thesis, Technische Universität Wien, 2005.
[31]A. J. Archer, Ph.D. thesis, University of Bristol, 2004.