Ground states of ultrasoft particles with attractions: a genetic algorithm approach

Christos N. Likos, Tobias Tueckmantel, Federica Lo Verso

Ground states of ultrasoft particles with attractions: a genetic algorithm approach

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

molphys.tgz

Ground states of ultrasoft particles with attractions:

a genetic algorithm approach

Tobias Tückmantel,[1] Federica Lo Verso,[2] and Christos N. Likos[1]

[1]*Institute for Theoretical Physics II: Soft Matter,*

Heinrich-Heine-Universität Düsseldorf,

Universitätsstraße 1, D-40225 Düsseldorf, Germany

[2]*Institut für Physik, Johannes-Gutenberg-Universität Mainz,*

Staudinger Weg 7, D-55099 Mainz, Germany

(Dated: January 28, 2009)

(A contribution to the special issue on the 65th birthday of Jean-Jacques Weis)

Abstract

We analyze in detail the ground-state structure of model systems of athermal star polymers with an additional, tunable attraction that may result from dispersion or depletion forces. To perform a free, unbiased search in the space spanned by the crystal parameters, we employ genetic algorithms, which are enhanced with respect to previous versions in their ability to find stable structures that occupy very narrow basins of attraction in the energy landscape. Application of this method brings about a very large variety of ground states for star polymers with attractions, in particular for the case of intermediate functionalities and strong, short-range attractive forces.

PACS numbers: 05.10.-a, 82.20.Wt, 64.70.Dv, 61.50.Ah

1

I. INTRODUCTION

Polymer-based, mesoscopically-sized colloidal particles are the subject of intensive research across a broad spectrum of disciplines in modern soft matter science. From chemical synthesis, to physical chemistry, scattering techniques, rheology, theory and simulation, the possibility to control the single-particle and collective properties of solutions of such colloids via modifications of their architecture, composition and solvent conditions, has attracted the sustained interest of scientists during at least the last two decades. A particularly intriguing subclass of polymer-based particles are those with a relatively low internal monomer concentration, a property that makes such particles soft and penetrable in comparison to conventional colloids. Evidently, the notion of *softness* is a relative one and it needs a proper, mathematical definition. A possibility is offered by the concept of the effective interaction $V_{\text{eff}}(r)$ between two colloidal particles [1–3]. The term *ultrasoft colloids* has been coined to characterize particles with an effective interaction that is *integrable*: $\int_0^\infty \text{d}^3r\, V_{\text{eff}}(r) < \infty$. There are many examples of polymer-based colloids whose effective interactions fulfill the above condition and which are therefore members of the class of ultrasoft colloids. As far as uncharged systems are concerned, star polymers [4–13] and dendrimers [14, 15] are prominent examples, whereas in the presence of charge, star-shaped polyelectrolytes [16–22] and ionic microgels [19, 23, 24] are ultrasoft colloids that have been studied extensively in the last decade.

All ultrasoft colloids share some common characteristics in their structural and phase behavior, such as the weakening of spatial correlation at high concentrations, re-entrant melting and even the complete absence of crystallization below some threshold values of polymer-arm number or charge [10, 24]. In addition, crystal structures stabilized by ultrasoft colloids tend to be open, i.e., they have a low coordination number, and hard to guess on the basis of intuition alone. These facts have been established for *purely repulsive* ultrasoft interactions, which are indeed emerging for all polymer systems when the solvent is athermal [1] or when the polymers carry charge. However, there are many occasions for which additional attractions are superimposed on an ultrasoft repulsion and their effects have only started to be explored in the recent years. The most straightforward mechanism leading to additional attractions between star polymers is to worsen the solvent quality, an effect that also leads to shrinkage of the stars and has quite intriguing consequences on the

2

dynamical arrest of the same [25].

A broad range of possibilities to induce *effective* attractions in athermal star polymer solutions is opened up by considering the ability to add smaller homopolymer chains or stars in the solution. The latter act then as depletants for the big star polymers and attractive contributions of varying strength and range can result, depending on the type and concentration of the depletant. The depletion attractions due to smaller stars and linear chains have been analyzed quantitatively by Mayer *et al.* [26, 27] using numerical simulations and scaling theory. The effects of the additives on the structural arrest in athermal star polymers are manifold and have been examined for both cases of added linear [28, 29] and star [30, 31] homopolymer. Free homopolymer chains that induce attractions to the stars can be found in solution as a result of the technique employed to synthesize the stars, in particular when the latter are formed *via* self-organized micellization of diblock copolymers [32]. The effects of the linear chains are then visible both in the correlations of the amorphous, fluid phases [33, 34] and in the stability of allotropic crystals at high concentrations [35–37]. Finally, another possibility to induce attractions is to graft the polymer chains on sizable colloidal particles and tune the dispersion attractions via changes of the temperature or the Hamaker constant [38].

Systematic efforts to analyze the effects of attractions on static correlations between star polymers and on the dynamical arrest of the same has been undertaken by Lo Verso *et al.* [39–42], who considered a specific functional form of the attractive contribution having the shape of a Fermi-Dirac distribution. It has been found that changing the depth and range of the attraction, combined with the number of arms (functionalities) f of the stars gives rise to various different thermodynamic scenarios, including the emergence of two critical points [39], and allows us to shift the glass lines of the system [40, 42]. Moreover, for low functionalities, estimates of the phase behavior of star polymers with additional attractions that are based on the Hansen-Verlet criterion have led to the conclusion that the attractions can induce both re-entrant crystallization [41] and re-entrant vitrification [42] that are both *absent* for purely repulsive, ultrasoft interactions. Nothing is known, however, on the nature of the crystal structures that are stabilized by attractive interactions when the latter are superimposed on ultrasoft repulsions. The purpose of this work is to fill this gap and provide a systematic analysis of the ground state phase diagram of star polymers with additional attractions, whereby we employ for the latter the model introduced in Ref. [39] and further

3

investigated in Refs. [40–42]. Recent results [43] have shown that the model introduced by Lo Verso *et al.* [39] is not an academic example but it rather can be realized in mixtures of star polymers and homopolymer linear chains.

The rest of this work is organized as follows: In section II we present the method we employed to trace out the phase diagram. It is based on genetic algorithms, which have been enriched in this work with additional features that enhance their performance in exploring efficiently the search space. In section III we present our model interaction potentials, with and without attractions, and in section IV the results and discussion of the phase diagrams. Finally, in section V we summarize and present our conclusions.

II. ENHANCED GENETIC ALGORITHMS

In this section we first present a concise summary of the ideas behind Genetic Algorithms in general and subsequently we discuss in detail the technical and algorithmic improvements we implemented with the goals of reducing the computational cost and of sampling regions of parameter space that might escape in the original version of the method.

A. Genetic Algorithms in general

Whereas steeply diverging interactions lead to crystals whose lattice structure is determined mainly by packing considerations, for soft interactions things are much less trivial. A variety of unusual, 'exotic' structures can be stabilized even by simple, spherically symmetric ultrasoft pair potentials [10, 44–46]. In this sense, simple guessing of the possible crystal structures is of no help for ultrasoft interactions, as almost certainly some competitive ones will be missed because they are difficult to guess. One needs an unbiased search strategy within the parameter space to find those crystal structures that minimize, e.g., the $T = 0$ free energy (i.e., the internal energy U of the system), which can then also be employed as candidates for a $T \neq 0$ calculation of the free energies. A very efficient tool for that purpose is offered by genetic algorithms (GA) [23, 24, 47], whose logic and implementation is shortly described below.

GAs can be considered as optimization strategies that use features of evolutionary processes as key elements; their purpose is to find optimal solutions for a given problem [48].

4

Originally, they have been developed by Holland and co-workers in the context of engineering science [49] and applied in many fields [48], such as economics, immunology, biology, or computer sciences. The pioneering work of Deaven and Ho [50] played a pivotal role in introducing GAs into Physics with the specific purpose of geometry optimization for given interactions. Recently, these tools have been extended to the determination of crystal structures for *infinite* systems in the bulk [24, 51–55] and in confined geometries [56].

The structural unit of a GA is an individual $\mathcal{I}$ which, in turn, is built up by a fixed number of 'genes'. In problems of 3d-crystallization, this individual encodes the three lattice vectors of the fundamental unit cell as well as the vectors of the basis, if any. In applications of finite cluster formation, all vectors of the cluster build up the corresponding individual. Depending on application, tradition and taste, one may choose to represent the position vectors, which are variational quantities, either as real (decimal) numbers or in a binary representation, i.e., as a succession of 1's and 0's. In the latter case, which is the one we adopt for the problem at hand, suitable generalized coordinates q_i that describe the unit cell are normalized to values $0 \leq q_i \leq 1$ and are represented in the binary alphabet, for details we refer the reader to Ref. [47]. The number of genes (binary digits) devoted to each generalized coordinate q_i is fixed and by concatenating successions of genes after each other, an individual is built that uniquely describes a crystal structure. An exemplary form of an individual is shown in Figure 1.

As a second ingredient of the GA, each individual $\mathcal{I}$ is assigned a *fitness value* derived by evaluation of the fitness function $f(\mathcal{I})$. The form of the fitness function must be chosen in such a way that the lower the (free) energy $F(\mathcal{I})$ of the crystal represented by the individual $\mathcal{I}$ is, the higher the (positive) value of its fitness $f(\mathcal{I})$. Though the choice of the form of $f(\mathcal{I})$ is *not* unique, we have empirically found that a suitable parametrization of the fitness is given by:

$$f(\mathcal{I}) = \exp\{-[F(\mathcal{I}) - F(\mathcal{I}_{\text{fcc}})]/F(\mathcal{I}_{\text{fcc}})\}; \tag{1}$$

where $F(\mathcal{I})$ is the free energy for a crystal structure represented by the individual $\mathcal{I}$ and $\mathcal{I}_{\text{fcc}}$ is the individual representing the fcc lattice; division of the exponent by $F(\mathcal{I}_{\text{fcc}})$ renders all values of the former to be of order unity, avoiding in this way the technical difficulties associated with exponentiating very large numbers. The free energy of any given crystal structure can be calculated, e.g., within a harmonic theory in the approximation of the Einstein model, or one can employ other approaches, such as phonon theory or density

5

functional theory. For the calculation of the ground states, $F(\mathcal{I})$ reduces to the lattice sum, i.e., the internal, interaction energy per particle in the crystal. We note that the evaluation of $F(\mathcal{I})$ is the most severe bottleneck in this approach, thus economical ways of evaluating the free energy of crystals are crucial for convergence in reasonable amounts of time.

A GA starts with a large number of individuals, typically a few hundred to a few thousand ones, which are generated randomly, i.e., as random successions of 1's and 0's. These form the first *generation* of the population, so that each individual $\mathcal{I}$ can be assigned a double index, $\mathcal{I}_j^{(\alpha)}$, Latin subscripts denoting the individual *within* the generation and Greek superscripts the generation itself. Thus, we begin with a collection $\{\mathcal{I}_j^{(1)}\}$ $j = 1, 2, \ldots, j_{\max}$. The next step is to choose two individuals as parents of two children for the next generation. The probability with which one individual is chosen as a parent is *proportional* to its fitness value. The combination between the mother's and the father's genes is performed by cutting the two individuals at a randomly chosen place (the same for the two parents but different for different parent pairs) and combining the first part of the mother with the second of the father to create the first child and vice versa to create the second one. This *crossover operation* is performed $j_{\max}/2$ times in each generation, creating thereby $j_{\max}$ children that form the next generation. The procedure is performed repeatedly until, after a number of generations (typically, a few hundred ones) the population has become fairly homogeneous and the fitness of the best individual of the generation does not improve any more. This fittest individual of the last generation is then taken as the optimal individual (crystal lattice) for the problem at hand.

Encoding of structures, evaluation of fitness, choice of the recombination probability and the recombination algorithm are, thus, four fundamental ingredients of the GA. A fifth, crucial one, is mutations. Guided by the incessant mutations taking place in living organisms, and which play a key role in evolution and adaptation, we also introduce a mutation probability p_m, typically of the order of 1% into the algorithm to avoid complete and premature homogenization of the population towards a sub-optimal individual. To this end, we randomly flip with probability p_m in a given generation every gene from 1 to 0 or vice versa. In this way, we introduce fresh genetic material into the pool and "thermalize" the system, helping to drive it out of local, metastable minima. It should be emphasized at this point that there is no guarantee that a GA will definitely converge to the absolute minimum; several runs of the GA should be performed at any rate and if these all converge

6

to the same minimum, one gains confidence that the latter is indeed the best one. A decisive advantage of the GAs in comparison with other search techniques is that they can perform arbitrarily large 'jumps' in the multidimensional search space through the recombination and mutation processes and, in this way, they are capable of escaping local minima in an efficient fashion.

Instead of just determining crystal structures for selected values of, e.g., density ρ or functionality f of the star polymers (see section III below), we wish to determine the whole phase diagram in the (ρ, f)-plane. This requirement puts high computational burdens in terms of speed, efficiency and accuracy, and requires improvements over the hitherto applied GA-techniques, which we will discuss in some detail in sections II B and II C. In addition to those, one needs to introduce and implement criteria that allow for the determination of phase boundaries, i.e., algorithms that can recognize a phase transition from one crystal structure to another as some boundary in the (ρ, f)-plane is crossed. For this purpose, we have implemented an edge-detection procedure capable of spotting first order phase transitions. Let there be two intervals of certain variables, e.g., ρ and f, which serve as parameters for determining optimal crystal structures. Then, first of all, we create a mesh on the Cartesian product $[\rho_{\min}, \rho_{\max}] \times [f_{\min}, f_{\max}]$. Second, we start a cyclic algorithm of dynamic mesh refinement: If the crystal structures determined for two adjacent parameter points differ by more than a predefined value, the associated rectangle is refined, i.e., it is divided into four equal smaller rectangles. This procedure is being repeated until some kind of accuracy condition is fulfilled, and thus it focuses on discontinuities of lattice values.

B. Speedup of the algorithm

We would like the genetic algorithm to become a reliable tool for automatically generating phase diagrams and discovering new, exotic crystal structures at the same time. If only using it on a few parameter sets, one will not encounter any runtime problems, since the execution of the GA we have implemented is a matter of minutes. The situation changes if one wants to generate a whole phase diagram, includes the subgeneration method and, hoping to find new structures, increases the number of individuals. While the phase border detection algorithm described above is running, it will start the GA several tens of thousands of times. Thus, there is a great need for speed improvements. Since the main computational effort lies in

7

the calculation of lattice sums, we will take a closer look upon that part of the program and propose a way to achieve significant speed improvement of the code.

Consider a Bravais lattice spanned by the elementary vectors $\vec{a}_1$, $\vec{a}_2$, and $\vec{a}_3$ and consisting of all points $\vec{R} = h\vec{a}_1 + k\vec{a}_2 + l\vec{a}_3$, where $[h, k, l] \in \mathbb{Z}^3$. One naive approach for computing a lattice sum would be to do a loop over all lattice points in a parallelepiped having the edge lengths $2h_{\max}$, $2k_{\max}$ and $2l_{\max}$, and containing the cutoff sphere:

$$M_{\text{cut}}(i, j) := \{[h, k, l] \in \mathbb{Z}^3 : |(h + d_{ij,1})\vec{a}_1 + (k + d_{ij,2})\vec{a}_2 + (l + d_{ij,3})\vec{a}_3| < r_{\text{cut}}\}, \quad (2)$$

with $\vec{d}_{ij} = \vec{b}_i - \vec{b}_j, i, j \in \{1, \ldots, N_b\}$ being the shift of the lattice associated with two different of the N_b basis particles located at $\vec{b}_i$ and $\vec{b}_j$, and $d_{ij,k}$ being the component of $\vec{d}_{ij}$ along the direction of the vector $\vec{a}_k$, $k = 1, 2, 3$. In this way we calculate the sum of the potential function evaluated at every point within the set M_{cut}, where r_{cut} is a cutoff parameter that determines the number of lattice points to be taken into account and is determined by a pre-determined convergence condition of the lattice sum. To avoid the expensive computation of square roots or the interaction potential, the latter is already tabulated in lookup tables as a function of the squared distance r^2. This method is quite fast and, in addition, necessary for the algorithm that follows. However, if the minimum distance $r_{\min}$ occurring in the lattice is very small, one should be extremely cautious due to the diverging discretization error of the interaction potential $V(r)$ as $r \to 0$. In these cases, it its preferable to evaluate the function directly at these positions.

The algebraic operations necessary to get a lattice vector for a certain set of indices $[hkl]$ are one three-dimensional matrix-vector product and an inner product. Thus, assuming $h_{\max} = k_{\max} = l_{\max} = N$ to simplify matters, one needs $O(N^3)$ of these operations. More-over, typecasts from a floating point number to an integer number in C++ take quite a long time, so it appears advisable to avoid them. For this reasons, we now suggest a modified algorithm, which is given in pseudo-code in the Appendix. Its principle is to perform first a cycle over h and k and compute the position vectors $\vec{r}_{hk0} = h\vec{a}_1 + k\vec{a}_2 + \vec{d}_{ij}$ corresponding to the index $[hk0]$ for every h, k. Furthermore, we evaluate the square of its norm, twice its inner product with the third primitive unit vector, $2\vec{r}_{hk0} \cdot \vec{a}_3$, and the squared norm $\vec{a}_3^2$. From now on, in this cycle, we stop using floating point representations of the squared distances; instead, all computations are done just with integer numbers, namely the indices corresponding to certain values in the look-up table. In the innermost loop, we make use of

8

the fact that for given vectors $\vec{r}, \vec{u} \in \mathbb{R}^n$ it holds:

$$|\vec{r} + l\vec{u}|^2 = (\vec{r} + (l-1)\vec{u})^2 + 2\vec{r} \cdot \vec{u} + 2(l-1)\vec{u} \cdot \vec{u} + \vec{u}^2$$
$$= |\vec{r} + (l-1)\vec{u}|^2 + 2\vec{r} \cdot \vec{u} + (2l-1)\vec{u}^2. \tag{3}$$

Thus, it is possible to compute the squared distance of a lattice point $[hkl]$ out of the squared distance of its predecessor. Since inner products $\vec{r} \cdot \vec{u}$ and $\vec{u}^2$ are constant in such a loop, one can reduce the number of matrix-vector products from $O(N^3)$ to $O(N^2)$, while the innermost iteration only contains two integer additions. In addition, one has to make sure this loop starts from the position of minimum distance, which, in the case of non-zero basis shifts, is not necessarily $[hk0]$. Instead, we choose the starting index l_{min} such that the length squared:

$$|\vec{r}_{hkl}|^2 = |\vec{r}_{hk0} + l\vec{a}_3|^2 = \vec{r}_{hk0}^2 + 2l\vec{r}_{hk0} \cdot \vec{a}_3 + l^2\vec{a}_3^2 \tag{4}$$

is minimized. One can easily see that

$$l_{min} = -\frac{\vec{r}_{hk} \cdot \vec{a}_3}{\vec{a}_3^2}. \tag{5}$$

If the lattice point $[h, k, l_{min}]$ is not located within the cutoff set $M_{cut}(i,j)$, one can be sure that no other point with same h, k is to be found there either.

This method is well-suited for very rapid computations of lattice sums L for a genetic algorithm; it generates discretization errors which result in a relative overall error $(L_{new} - L_{old})/L_{old} \approx 10^{-3}$, while working ten times as fast as the simple algorithm. We considered this to be sufficient for the requirements of a probabilistic optimization algorithm, provided one uses a non-probabilistic minimizer, such as a quasi-Newton, steepest-descent or conjugate gradient method towards a more precise value of lattice sum at the end [47].

C. Memorizing best individuals, subgenerations

During thorough tests of the GA, we encountered several problems related to the topography of the potential energy landscape and causing loss of information. Let us assume that the function $\mathcal{G}$ one would like to minimize has a local minimum with a broad basin of attraction, i.e., a large set of starting values would drive a classical optimization algorithm towards that minimum. Besides this broad minimum, $\mathcal{G}$ has got another one which is much lower, but very narrow and therefore difficult to hit.

9

In such cases the GA is very likely to find the broad minimum while failing to discover the narrow one, because genetic information coding the latter does not appear frequently enough by means of recombination or mutation. Thus, the recombination probability of these rare individuals is much too low and their information tends to get lost. We address this problem by introducing two mechanisms of process control: first, the storage of best individuals, and second, the use of subgenerations.

The first method means to store the individual of best fitness value in each generation step, an approach known as *elitism*. In this manner one creates an additional FIFO-buffer of best individuals which reaches into the past and contains n_V parameter sets. When this buffer is full, the oldest individual is dropped and replaced by a new one. From this moment, every new generation only consists of $n_G - n_V$ parameter sets created by recombination while the remaining n_V individuals are being copied from the buffer. In doing so, we ensure that essential data gained throughout the process does not get lost that easily. Typically, we are working with a best-individual list containing $n_V = 50$ individuals, whereas each generation consists of $n_G = 2000$ individuals in total.

Using this method, we have been able to greatly improve the performance of the GA; yet, the drawback of narrow-minimum individuals being suppressed has not been removed, because of a second mechanism taking place: if certain individuals are actually located in such a critical area but do not hit its bottom precisely enough, their lattice sum will exceed that of less promising parameter sets. In such cases, even the method mentioned above does not help at all and we need an additional refinement of our technique. Our suggestion is to use subgenerations, which means to apply a convergence criterion in order to decide whether the GA appears to be aiming for a certain location. Then the program will split the population into two equal subgenerations, which differ in the way they are filled with new individuals by means of recombination. The first of these parts is just being processed as the whole generation has been before, but the second one is treated differently: when one needs to calculate fitness values in order to get recombination probabilities for the new individuals of subgeneration 2, we apply a scaling factor $s(d(\mathcal{I}, \mathcal{I}_1)) \in [0, 1]$, where $\mathcal{I}$ is the individual whose fitness we are computing and $\mathcal{I}_1$ is the best parameter set of subgeneration 1. With $d(\mathcal{I}, \mathcal{J})$ we denote a distance function expressing how different two individuals are, while $s : [0, \infty[\rightarrow [0, 1]$ is chosen to be monotonic increasing with $s(0) = 0$. Thus, we make the second subgeneration avoid the best parameter set of the first one. However, we still use the

10

entire set of individuals as a basis for recombination. This method is not restricted to a total of two subgenerations. If three or more subgenerations are used, the procedure described above just has to be generalized in order to let the individuals of the j-th subgeneration avoid the best constructs of those with indices $\{1, \ldots, j-1\}$.

III. MODEL INTERACTION POTENTIALS

A. Purely repulsive pair potential

The reference physical system are star polymers in athermal solvents, for which case the effective interaction potential between the centers of the stars is purely repulsive. As the coarse-grained potential arises from microscopic excluded-volume interactions, it is *entropic* in origin and concomitantly *athermal*, i.e., it scales proportionally to the thermal energy $k_\mathrm{B}T$, with Boltzmann's constant k_B. Its functional form depends parametrically on the number of arms (functionality) f of the stars and reads as [9]:

$$\beta V_\mathrm{rep}(r) = \frac{5}{18} f^{3/2} \begin{cases} \left[-\ln\left(\frac{r}{\sigma}\right) + \frac{1}{1+\sqrt{f}/2} \right] & \text{if } r \leq \sigma; \\ \frac{1}{1+\sqrt{f}/2}\left(\frac{\sigma}{r}\right) \exp\left[-\frac{\sqrt{f}(r-\sigma)}{2\sigma} \right] & \text{if } r \geq \sigma, \end{cases} \tag{6}$$

where $\beta = (k_\mathrm{B}T)^{-1}$ and σ is the corona diameter of the stars, related to the experimentally measurable radius of gyration R_g via $\sigma \cong 1.26\, R_g$ [57]. The phase diagram of athermal star polymers, drawn in the plane spanned by the functionality f and the density ρ of stars has been investigated in Ref. [10]. There, the competition between fluids and crystals has been considered at finite temperature T. Motivated by the unusual crystal phases appearing there, the ground states of the system have been examined with the help of genetic algorithms for a single functionality $f = 64$ by Gottwald *et al.* [47]. Here, we revisit the system as a point of reference and we improve on the work of Ref. [47] on two counts: first, in that we employ the enhanced version of the algorithm sketched in section II above and second in that we now draw the ground-state phase diagram for an extended region in the (ρ, f)-plane. We will indeed see that the enhanced algorithms find phases that the original version had missed.

11

B. Attractive contributions

On top of the repulsion given by Eq. (6), we now superimpose an attraction $w(r)$, introducing thereby the total interaction potential:

$$V(r) = V_{\text{rep}}(r) + w(r). \tag{7}$$

The form of $w(r)$ follows from the physical motivation put forward in Refs. [39, 40]. In order to maintain the $\sim 1/r$-scaling of the effective force $F(r) = -\nabla V_{\text{tot}}(r)$ at short distances, the attractive contribution $w(r)$ is chosen to have the form of a Fermi potential, which is essentially constant for small r-values, and whose point of inflection and sharpness are free parameters. In particular, $w(r)$ reads as:

$$w(r) = -C \left\{ \exp\left[\frac{r - A}{B} \right] + 1 \right\}^{-1}. \tag{8}$$

Let us briefly discuss the meaning of the various parameters and the physical mechanisms that allow them to be externally controlled. C is an energy scale, whereas A and B are length scales. The relevant physical quantity is the reduced attraction strength,

$$C^* \equiv \frac{C}{k_{\text{B}}T}. \tag{9}$$

If the attraction results from dispersion forces, then C can be tuned by modifying the solvent that affects the value of the Hamaker constant [38]. In this case, C is temperature-independent and the system is thermal. If the attraction is caused due to the addition of athermal depletants (such as polymer coils [27] or smaller stars [26]), then C scales with $k_{\text{B}}T$, hence temperature is irrelevant, but it acquires a dependence on the depletant concentration. In this case, the above-defined quantity C^* can be steered by changing the depletant reservoir concentration. Attractions can also be induced by changing the solvent quality [58], for which case C^* can be roughly associated with the quantity T_Θ/T, where T_Θ is the Θ-temperature of the polymer and in the limit $T \gg T_\Theta$ we obtain the athermal solvent conditions, $C^* = 0$.

The length scales A and B can also be tuned externally. In the case of dispersion-induced attraction, they can be influenced by screening the dipole-dipole interaction through addition of salt, which remains mainly in the intermolecular space, not affecting therefore the conformation of the stars. In the case of depletant-induced attraction they can be influenced

12

by modifying the softness and size of the smaller component. In what follows, we make the choice $A = 1.35\,\sigma$ and $B = 0.05\,\sigma$, where σ is the star size that sets the length scale of the purely repulsive interaction, Eq. (6). Indeed, for this parameter combination, preceding work has established the occurrence of intriguing self-organization phenomena, including an interplay between cluster formation and macrophase separation [41], as well as the existence of broad region of unknown crystalline structures and reentrant glasses [41, 42]. In Fig. 2 we show the resulting total interaction potential $V(r)$, Eq. (7), for the choice of the parameters A and B mentioned above, star functionality $f = 20$ and different values of the attraction strength C^*.

IV.   RESULTS

A.   Purely repulsive star polymer solution

We first revisited the problem of the ground-state phase diagram of purely repulsive star polymers, i.e., ultrasoft particles interacting by means of the potential given in Eq. (6), as a means of testing the novel, enhanced version of the genetic algorithm. In Ref. [47], Gottwald et al. have already applied the GA-method to determine the ground-state structures for a single value of the functionality, $f = 64$ in the density range $0 \leq \rho\sigma^3 \leq 4$. Here, we expand drastically by considering densities in the range $0.5 \leq \rho\sigma^3 \leq 5.0$ and functionalities $20 \leq f \leq 200$. $f \in \{20..200\}$.

The resulting phase diagram is shown in Fig. 3. For the case $f = 64$ and below the density $\rho\sigma^3 = 4.0$, the previous results of Ref. [47] are confirmed, with one important exception, namely: in the density range $2.6 \lesssim \rho\sigma^3 \lesssim 3.7$, part of the region of stability of the hexagonal crystal is replaced by a new crystal that has not been seen before. It is a general, triclinic lattice with three basis particles per unit cell, a structure we call S1 and whose unit cell is shown in Fig. 4. This is one example of the benefits we obtain by introducing the subgenerations, since indeed the S1-structure (and others to follow) are typical examples of deep but narrow minima in configuration space. For densities of $1.3 < \rho\sigma^3 < 2$, we observe triclinic lattices with some noisy spots, an artifact of the algorithm for automatic generation of phase boundaries, caused by two coexisting phases. To additionally check our results, we have computed for the most asymmetric phases the magnitude Λ of the total force acting

13

on every basis particle due to all others:

$$\Lambda = \sum_{j=1}^{n} |\vec{F}_j| = \sum_{j=1}^{n} | \sum_{k \in \{1..N\} \setminus \{n\}} \sum_{\vec{R}_i \in \mathcal{G}} \vec{F}(\vec{d}_n - (\vec{R}_i + \vec{d}_k)) + \sum_{\vec{R}_i \in \mathcal{G} \setminus \{0\}} \vec{F}(-\vec{R}_i)|. \qquad (10)$$

For an equilibrium structure, one should get $\Lambda = 0$. Indeed, the value we obtain for Λ is vanishingly small after we apply a steepest-descent minimization on top of the structure predicted by the genetic algorithm, as usual [47], a process that leads to the final minimum. This thus offers an additional confirmation for the stability of the S1- and other novel phases discovered by the enhanced genetic algorithms.

The region located in the lower right corner of the diagram of Fig. 3 features the A15 structure, already suggested in Refs. [44, 45] for this range. However, the wedge-shaped, S2-region above the A15 in Fig. 3 has not been predicted before. Instead of a continued border between the A15 and the hcp-phases as f grows, the new GAs discover a new crystal, denoted by S2, which is similar to the previously discussed S1-phase: it is, like the former, a triclinic Bravais lattice with three basis particles, and a typical configuration of its unit cell is shown in Fig. 5. Moreover, the system behaves again unexpectedly in a small 'pocket' around when $f \approx 55$ and $\rho\sigma^3 \approx 3$. There, a structure which we call S5, whose unit cells are drawn in Fig. 6, shows up. It is a lattice with two basis particles, for which one angle between two of the edges is close to $\pi/3$. We emphasize that the three new structures found thus far, the S1, the S2 and the S5 are *all* the result of the enhancements of the genetic algorithms.

B. Star polymers with attractive interaction

We now switch to the full potential, $V(r) = V_{\text{rep}}(r) + w(r)$ given by Eqs. (6)-(8). As mentioned above, we employ the length scales $A = 1.35\sigma$, $B = 0.05\sigma$ and increase the attraction strength gradually, from $C^* = 0$ (preceding section) to the values $C^* = 1, 2$ and 5.

The ground state phase for the case $C^* = 1$ is shown in Fig. 7. Comparing with the purely repulsive case $C^* = 0$ in Fig. 3, we see that the succession of phases and even the phase boundaries remain largely unaffected for the high f-values, at the upper part of the phase diagram. This is to be expected since, for $f \gg 1$ the attractive contribution $w(r)$ is a small perturbation to the reference, repulsive potential $V_{\text{rep}}(r)$. The effects of

14

the attraction are nevertheless clearly visible for functionalities $f \lesssim 100$ and at density values around $\rho\sigma^3 \cong 1.0$. Indeed, in this region we see in Fig. 3 the emergence of a bcc and a hexagonal phase that were absent for $C^* = 0$. The region of density for which this occurs can be rationalized by considering the typical nearest-neighbor distance a_{fcc} of the fcc-lattice at such densities, in conjunction with the form of the potential shown in Fig. 2. Using $a_{\text{fcc}} = (4/\rho)^{1/3}/\sqrt{2}$, we obtain $a_{\text{fcc}} \cong 1.12$. This is the point in which the potential displayed in Fig. 2 starts showing significant deviations from its form for $C^* = 0$ and lies close to a local minimum that $V(r)$ even develops for higher values of C^*. Accordingly, it is precisely at such densities that the system finds it more preferable to form a lattice (bcc) that has fewer nearest neighbors (eight) *closer* to a given lattice point than more (twelve) farther away from it; evidently, higher-order neighbors also play an increasingly important role as density grows, bringing about a new region of stability of the hexagonal lattice adjacent to the bcc in Fig. 7.

The algorithm employed to determine phase boundaries is designed to detect sufficiently abrupt changes in the lattice parameters within a 'pixel' $(\Delta\rho, \Delta f)$ of the phase diagram. As such, it can only find first-order phase transitions between sufficiently 'different' crystal structures but not boundaries in which one phase smoothly evolves into a different one. One such characteristic scenario develops for the case at hand around the region $f \approx 118$ and $1.1 \lesssim \rho\sigma^3 \lesssim 2.0$ in Fig. 7. Here, the triclinic phase that is stable at higher f-values evolves into a bcc-phase on the left and a hexagonal phase (hex) on the right. These and other second-order phase boundaries are hand-drawn in Fig. 7 and subsequent cases as straight blue lines. Additional changes due to the attractive contribution with $C^* = 1$ as opposed to the $C^* = 0$-case develop at the high-density part of the phase diagram, where the previous region of stability of the A15-phase is reduced to give place to a new, high-density bcc-crystal, see Fig. 7. Otherwise, we still find the 'wedge' of the S2-phase with slightly distorted phase boundaries compared to the purely repulsive case.

Raising now the strength of the attractive part to $C^* = 2$ results into the diagram shown in Fig. 8. To a significant extent we find a continuation of the trends already seen in going from $C^* = 0$ to $C^* = 1$. The bcc- and hex-phases around the region $1.0 \lesssim \rho\sigma^3 \lesssim 2.0$ extend their domains of stability at even higher f-values, whereas new phases start appearing at the bottom of the phase diagram and for functionalities $20 \leq f \lesssim 40$, forming small 'pockets' that are nested within each other. We will identify these pockets for the case of stronger

15

attractions, $C^* = 5$, below. More importantly, however, at the high-density part of the phase diagram, *two* new phases, S3 and S4 appear, which squeeze the previously stable S2 phase into a small region and push the A15-phase out of sight. Some unit cells of the S3- and S4-phase are shown in Figs. 9 and 10, respectively. The former is a triclinic crystal with four particles per unit cell, whereas the latter possesses a five-particle basis, i.e., they are both highly complex structures yet stabilized by spherosymmetric interactions. The parameters of the S5-phase around $1.7 \lesssim \rho\sigma^3 \lesssim 2.0$ and $f \lesssim 40$ look very similar to those of a diamond structure; therefore the algorithm fails in detecting the phase boundary between S5 and the diamond region.

In order to get further insight into the influence of the attractive interactions on ultrasoft repulsions, we have finally computed a diagram for $C^* = 5$, which is displayed in Fig. 11. It contains many new and interesting structures, including the S3-, S4- and S5 crystals, found before. Moreover, for moderate densities, $\rho\sigma^3 \sim 1.0$ and functionalities $f \lesssim 60$, we obtain a cascade of phase transitions among phases enclosed in small 'pockets', already mentioned in the discussion of the case $C^* = 2$ above. A new lattice, labeled fcc2, appears, which exhibits the same angles and length ratios as an fcc Bravais lattice, but has a two-particle basis. At this strength of inter-molecule attraction, the interaction potential shows local minima and considerable substructure. Counterintuitive, low-symmetry structures with bases become likely; this is because the distances which appear during evaluation of the lattice sum function get scattered and distributed in the potential dip. Hence, structures with lower symmetry benefit from additional degrees of freedom regarding their configuration, implying a potential advantage compared to those with high symmetry. Other structures found in the narrow pockets are a monoclinic one with a two-particle basis, a triclinic and even a low-density A15 lattice. These findings demonstrate the subtleties in the interplay between short-range attractions and ultrasoft repulsions, in particular for the choice of the A- and B-parameters that lead to the hump in the potential shown in Fig. 2.

V.   CONCLUSIONS

In this work, we have presented novel developments on two aspects of modern computational science, the algorithmic and the physical one. On the first count, we have further developed the technique of genetic algorithms in such ways that an efficient, detailed and ac-

16

curate exploration of search space is achieved, with the goal of identifying the stable crystal structures for given interactions. The novel developments allow for an efficient calculation of phase diagrams in two-dimensional parameter space, as is usual, e.g., in one-component systems. From the physical point of view, we then applied this technique to various cases of star polymers with additional attractions, a system whose ground-state structure had been unexplored to-date. The enormous variety of crystal structures discovered is a testimony to the efficiency *and* the necessity of employing unbiased search strategies, such as genetic algorithms, in exploring the phase diagram of such systems. Indeed, in such a case mere intuition fails and the exceedingly intricate shapes of the resulting phase diagrams require techniques of the kind employed in this work. Our results are testable in star-linear polymer mixtures, in which case the chains might actually act as agents that prevent the system from getting trapped into a metastable, glassy state [28] accelerating thereby the nucleation route to periodic crystals. Current work is being focused on the phase diagram of soft, *anisotropic* particles [59, 60] as well as on the investigations of the structure of soft particles in confined slit pores of varying width and the transition from the two- to the three-dimensional freezing limit [61].

Acknowledgments

APPENDIX A: PSEUDOCODE ON SPEEDUP FOR THE LATTICE SUMS

Here we present a pseudocode that implements the speedup algorithm outlined in section II B. Note that here the quantity $\vec{r}_{hk0}$ is called v.

```
irmax = index(rcut)
for h = -hmax to hmax do
   for k = -kmax to kmax do
       get distance vector v = h*a1 + k*a2 + bij
```

17

```
compute squared distance r2 = v*v

get index of r2 in the look-up table : ir = index(r2)

compute inner product va = 2*v*a3 and its Index iva

compute inner product aa = a3*a3 and its Index iaa

j = ir

sum = sum + Phi[j]

m = iaa

while ( j < irmax )

    j = j + iva + m

    m = m + 2*iaa

    sum = sum + Phi[j]

m = iaa

j = ir

while ( j > -irmax )

    j = j - iva + m

    m = m + 2*iaa

    sum = sum + Phi[j]
```

18

[1] W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersions* (Cambridge University Press, 1989).

[2] C. N. Likos, Phys. Rep. **348**, 267 (2001).

[3] C. N. Likos, Soft Matter **2**, 478 (2006).

[4] M. Daoud and J. P. Cotton, *J. Phys. (France)* **43**, 531 (1982).

[5] P. A. Pincus and T. A. Witten, *Macromolecules* **19**, 2509 (1986).

[6] T. A. Witten, P. A. Pincus, and M. E. Cates, *Europhys. Lett.* **2**, 137 (1986).

[7] G. S. Grest, K. Kremer, and T. A. Witten, *Macromolecules* **20**, 1367 (1987).

[8] G. S. Grest, L. J. Fetters, J. S. Huang, and D. Richter, *Adv. Chem. Phys.* **XCIV**, 67 (1996).

[9] C. N. Likos, H. Löwen, M. Watzlawek, B. Abbas, O. Jucknischke, J. Allgaier, and D. Richter, Phys. Rev. Lett. **80**, 4450 (1998).

[10] M. Watzlawek, H. Löwen, and C. N. Likos, Phys. Rev. Lett. **82**, 5289 (1999).

[11] D. Vlassopoulos, G. Fytas, T. Pakula, and J. Roovers, *J Phys: Condens. Matter* **41**, R855 (2001).

[12] G. Foffi, F. Sciortino, P. Tartaglia, E. Zaccarelli, F. Lo Verso, L. Reatto, K. A. Dawson, and C. N. Likos, Phys. Rev. Lett. **90**, 238301 (2003).

[13] D. Vlassopoulos, J. Polym. Sci. Part B – Polym. Phys. **42**, 2931 (2004).

[14] M. Ballauff and C. N. Likos, Angew. Chem. Int. Ed. **43**, 2998 (2004).

[15] J. J. Freire, Soft Matter **4**, 2139 (2004).

[16] N. P. Shusharina and M. Rubinstein, Macromolecules **41**, 203 (2008).

[17] A. Jusufi, C. N. Likos, and H. Löwen, Phys. Rev. Lett. **88**, 018301 (2002).

[18] A. Jusufi, C. N. Likos, and H. Löwen, J. Chem. Phys. **116**, 11011 (2002).

[19] A. R. Denton, Phys. Rev. E **67**, 011804 (2003).

[20] H. Wang and A. R. Denton, Phys. Rev. E **70**, 041404 (2004).

[21] H. Wang and A. R. Denton, J. Chem. Phys. **123**, 244901 (2005).

[22] A. Jusufi, C. N. Likos, and M. Ballauff, Colloid Polym. Sci. **282**, 910 (2004).

[23] D. Gottwald, C. N. Likos, G. Kahl, and H. Löwen, Phys. Rev. Lett. **92**, 068301 (2004)

[24] D. Gottwald, C. N. Likos, G. Kahl, and H. Löwen, J. Chem. Phys. **122**, 074903 (2005).

[25] M. Kapnistos, D. Vlassopoulos, G. Fytas, K. Mortensen, G. Fleischer, and J. Roovers, Phys.

Rev. Lett. **85**, 4072 (2000).

[26] C. Mayer, C. N. Likos, and H. Löwen, Phys. Rev. E **70**, 041402 (2004).

[27] C. Mayer and C. N. Likos, Macromolecules **40**, 1196 (2007).

[28] E. Stiakakis, D. Vlassopoulos, C. N. Likos, J. Roovers, and G. Meier, Phys. Rev. Lett. **89**, 208302 (2002).

[29] E. Stiakakis, G. Petekidis, D. Vlassopoulos, C. N. Likos, H. Iatrou, N. Hadjichristidis, and J. Roovers, Europhys. Lett. **72**, 664 (2005).

[30] E. Zaccarelli, C. Mayer, A. Asteriadi, C. N. Likos, F. Sciortino, J. Roovers, H. Iatrou, N. Hadjichristidis, P. Tartaglia, H. Löwen, and D. Vlassopoulos, Phys. Rev. Lett. **95**, 268301 (2005).

[31] C. Mayer, E. Zaccarelli, E. Stiakakis, C. N. Likos, F. Sciortino, A. Munam, M. Gauthier, N. Hadjichristidis, H. Iatrou, P. Tartaglia, H. Löwen, and D. Vlassopoulos, Nature Mater. **7**, 780 (2008).

[32] J. Stellbrink, G. Rother, M. Laurati, R. Lund, L. Willner, and D. Richter, J. Phys.: Condens. Matter **16**, S3821 (2004).

[33] M. Laurati, J. Stellbrink, R. Lund, L. Willner, D. Richter, and E. Zaccarelli, Phys. Rev. Lett. **94**, 195504 (2005).

[34] M. Laurati, J. Stellbrink, R. Lund, L. Willner, E. Zaccarelli, and D. Richter, Phys. Rev. E **76**, 041503 (2007).

[35] J. Bang, T. P. Lodge, X. Wang, K. L. Brinker, and W. R. Burghardt, Phys. Rev. Lett. **89**, 215505 (2002).

[36] T. P. Lodge, J. Bang, M. J. Park, and K. Char, Phys. Rev. Lett. **92**, 145501 (2004).

[37] J. Bang and T. P. Lodge, Phys. Rev. Lett. **93**, 245701 (2004).

[38] P. N. Pusey, in *Les Houches, Session LI, Liquids, Freezing, and the Glass Transition*, ed. by J.-P. Hansen, D. Levesque, and J. Zinn-Justin (North Holland, Amsterdam, 1991).

[39] F. Lo Verso, M. Tau, and L. Reatto, J. Phys.: Condens. Matter **15**, 1505 (2003).

[40] F. Lo Verso, L. Reatto, G. Foffi, P. Tartaglia, and K. A. Dawson, Phys. Rev. E **70**, 061409 (2004).

[41] F. Lo Verso, C. N. Likos, and L. Reatto, Progr. Colloid Polym. Sci. **133**, 78 (2006).

[42] F. Lo Verso, C. N. Likos, C. Mayer, and L. Reatto, Mol. Phys. **104**, 3523 (2006).

[43] M. Camargo and C. N. Likos, in preparation.

20

[44] P. Ziherl and R. Kamien Phys. Rev. Lett. **85**, 3528 (2000).

[45] P. Ziherl and R. Kamien Phys. Rev. Lett. **85**, 3528 (2000).

[46] C. N. Likos, N. Hoffmann, H. Löwen, and A. A. Louis, J. Phys.: Condens. Matter **14**, 7681 (2002).

[47] D. Gottwald, G. Kahl, and C. N. Likos, J. Chem. Phys. **122**, 204503 (2005).

[48] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Mass., 1989.

[49] J. H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, 1975.

[50] D. M. Deaven and K. M. Ho, Phys. Rev. Lett. **75**, 288 (1975).

[51] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).

[52] J. Fornleitner, F. Lo Verso, G. Kahl, and C. N. Likos, Soft Matter **4**, 480 (2008).

[53] J. Fornleitner and G. Kahl, Europhys. Lett. **82**, 18001 (2008).

[54] G. Pauschenwein and G. Kahl, Soft Matter **4**, 1369 (2007).

[55] G. Pauschenwein and G. Kahl, J. Chem. Phys. **129**, 174107 (2008).

[56] J. Dobnikar, J. Fornleitner, and G. Kahl, J. Phys.: Condens. Matter **49**, 494220 (2008).

[57] A. Jusufi, M. Watzlawek, and H. Löwen, Macromolecules **32**, 4470 (1999).

[58] S. Huißmann, R. Blaak, and C. N. Likos, Macromolecules **xx**, yyy (2009).

[59] S. Prestipino and F. Saija, J. Chem. Phys. **126**, 194902 (2007).

[60] A. Nikoubashman and C. N. Likos, in preparation (2009).

[61] M. Kahn, J.-J. Weis, C. N. Likos, and G. Kahl, submitted (2009).

21

| 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | ••• | 0 | 1 | 0 | 0 | 1 |

$$\underbrace{}_{q_1} \quad \underbrace{}_{q_2} \quad \cdots \quad \underbrace{}_{q_n}$$
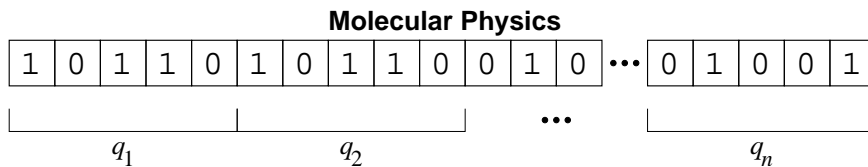
FIG. 1: A schematic drawing of an individual representing the unit cell and possible basis vectors of a crystal. Each generalized coordinate q_i, $i = 1, 2, \ldots, n$ is assigned a fixed number of genes.
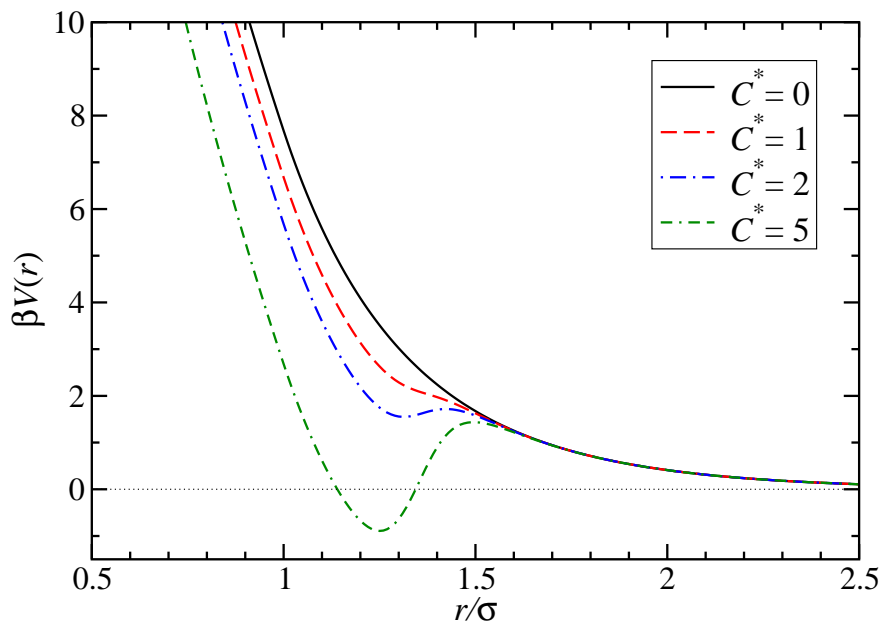

FIG. 2: The pair interaction potential $V(r)$ given by the sum of the purely repulsive term, Eq. (6), and the attractive part, Eq. (8). Results are shown for functionality $f = 20$ and various different values of the attraction strength C, as indicated in the legend.

22
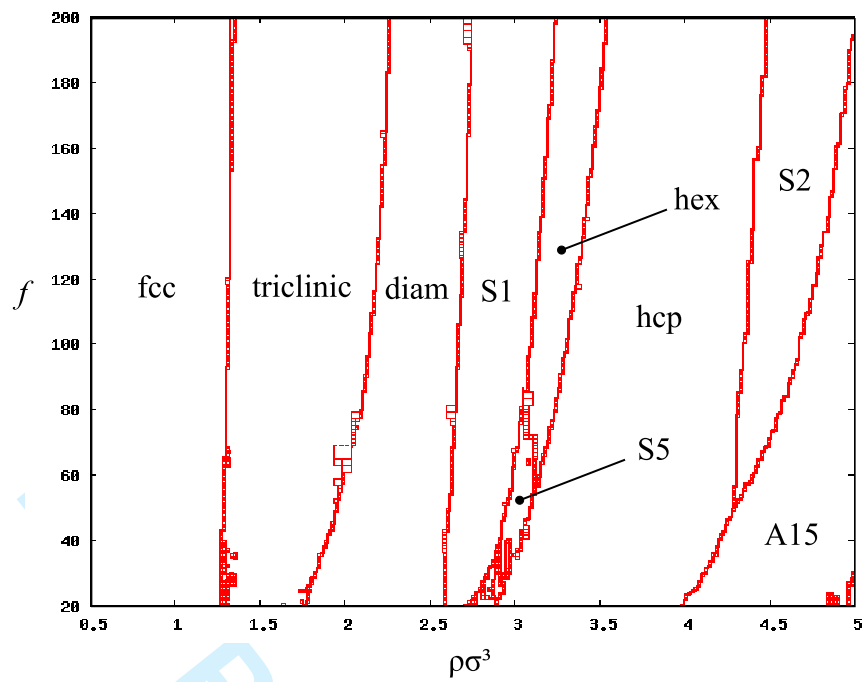
FIG. 3: The ground-state phase diagram for star polymers without additional attraction. The enhanced genetic algorithms bring about hitherto undiscovered crystals as stable ones, namely the new structures termed S1, S2 and S5. Unit cells of these are shown in Figs. 4, 5 and 6 that follow.
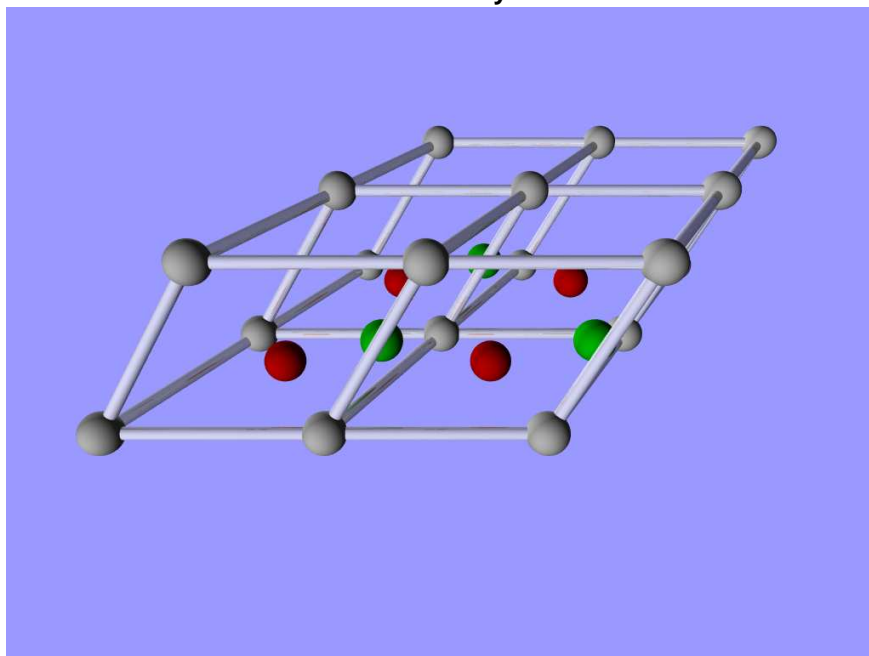
FIG. 4: Unit cells of the crystal structure S1 of Fig. 3 at a typical point, $f \approx 100$, $\rho\sigma^3 \approx 2.8$. The lattice- and basis-parameters of this phase change very little with f and ρ within its domain of stability.

24
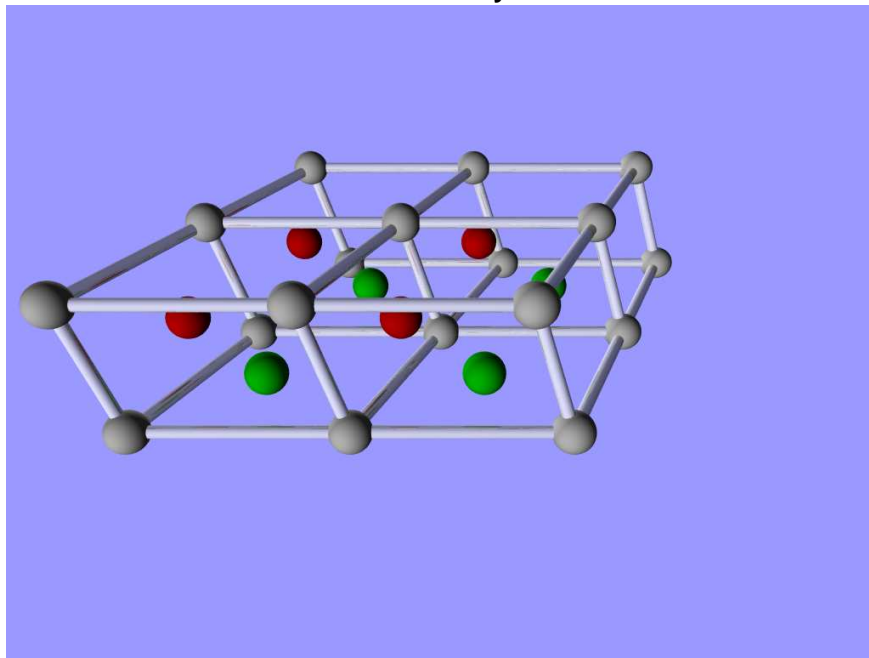
FIG. 5: Unit cells of the crystal structure S2 of Fig. 3. Here, we show a snapshot for the case of star polymers with attraction, $C^* = 1$, Fig. 7, where this structure appears again, and for parameters $f \approx 120$, $\rho\sigma^3 \approx 4.5$. Due to the high value of f, the weak attraction has a very small effect on the potential and the system is very similar to star polymers without attraction, thus the crystal parameters of the phase S2 at this point are very similar to those in Fig. 3.

25
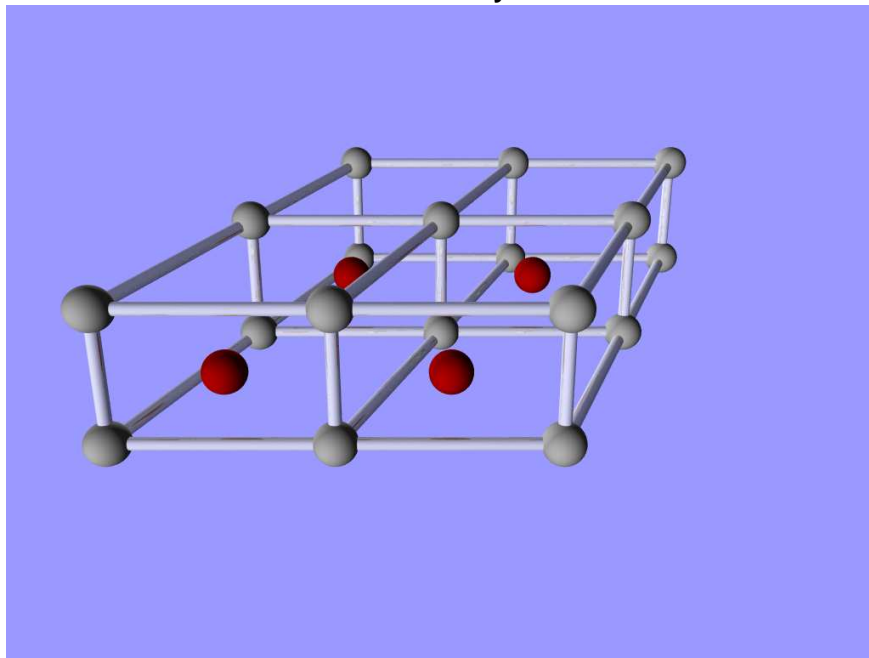
FIG. 6: Typical configuration of the unit cells of the phase S5, Fig. 3.
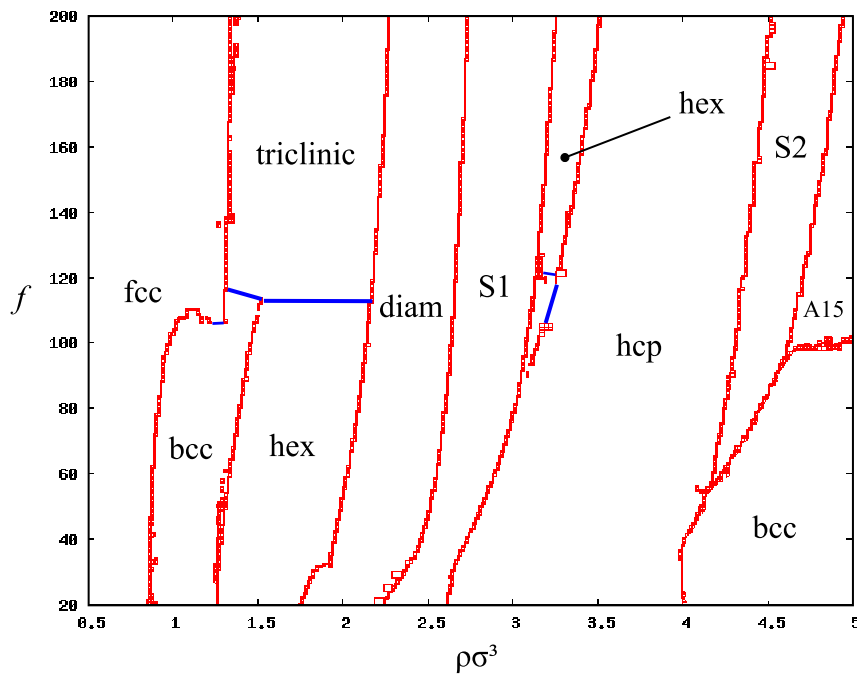
FIG. 7: The ground-state diagram for the star polymer system with an attractive contribution, Eq. (8), and for attraction strength $C^* = 1$.
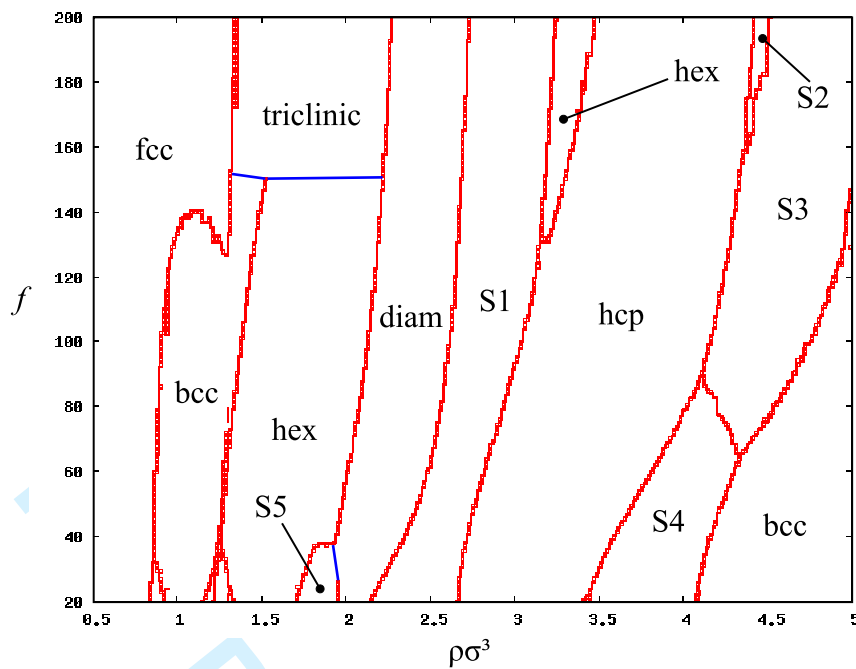
26

FIG. 8: Same as Fig. 7 but for $C^* = 2$. For a characterization of the phases appearing at the small 'pockets' around $\rho\sigma^3 \approx 1.0$ and $f \lesssim 40$, see Fig. 11.
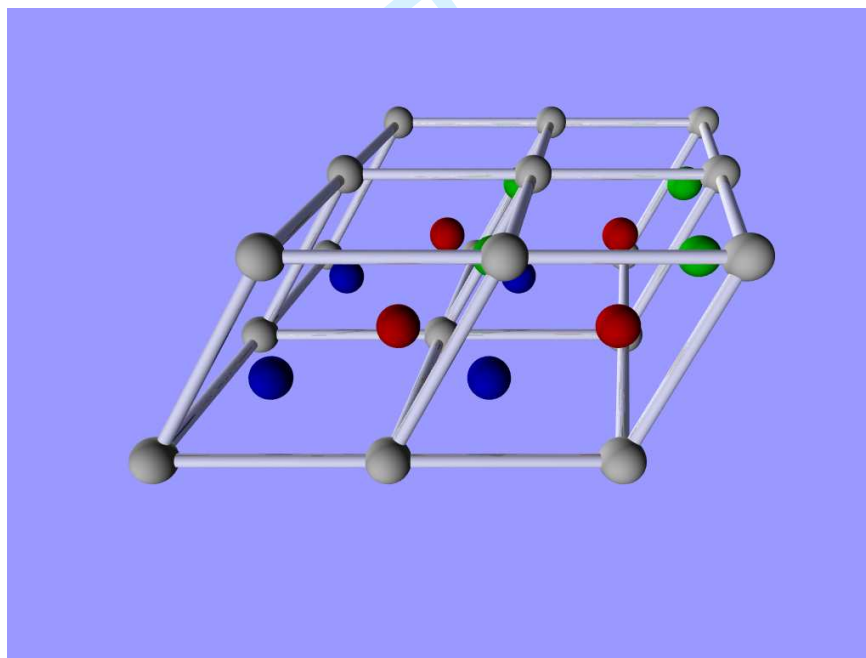

FIG. 9: Unit cells of the structure S3 found for the case $C^* = 2$, here at $\rho\sigma^3 \approx 4.5$, $f \approx 70$, see Fig. 8.
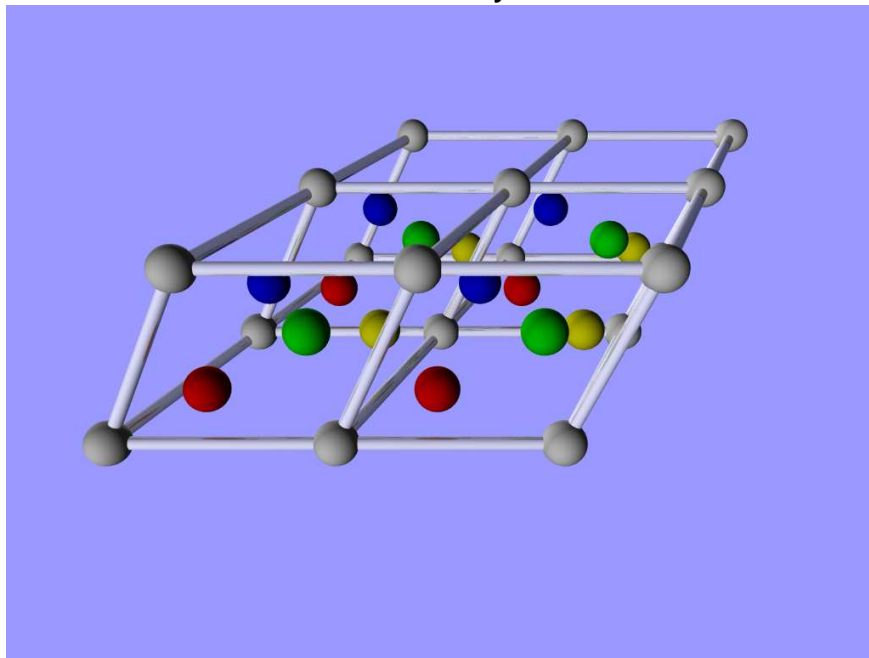
27

FIG. 10: Unit cells of the structure S4 found for the case $C^* = 2$, here at $\rho\sigma^3 \approx 3.0$, $f \approx 30$, see Fig. 8.
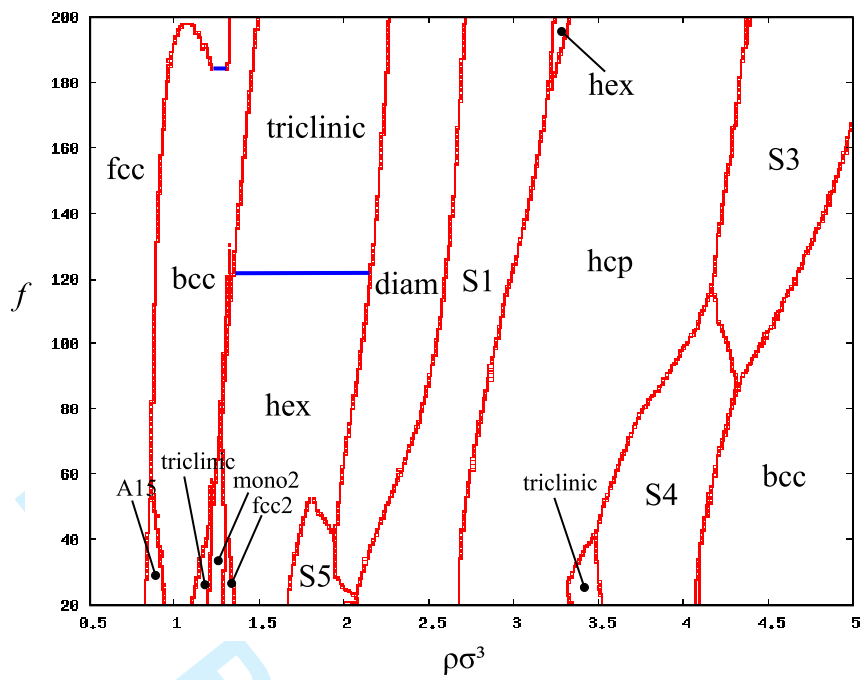
FIG. 11: Same as Fig. 7 but for $C^* = 5$. The region marked with fcc2 shows a structure very similar to an fcc bravais lattice and a two-particle basis; the name 'mono2' denotes a monoclinic lattice with a two-particle basis.

29