A new paradigm for structure prediction in multicomponent systems

D. Schebarchov[a)] and D. J. Wales[b)]

University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, UK

(Dated: 23 November 2013)

We analyse the combinatorial aspect of global optimisation for multicomponent systems, which involves searching for the optimal chemical ordering by permuting particles corresponding to different species. The overall composition is presumed fixed, and the geometry is relaxed after each permutation in order to relieve local strain. From ideas used to solve graph partitioning problems we devise a deterministic search scheme that outperforms (by orders of magnitude) conventional and self-guided basin-hopping global optimisation. The search is guided by the energy gain from either swapping particles i and j (ΔE_{ij}) or changing the identity of particles i (ΔE_i). These quantities are derived from the underlying (arbitrary) energy function, hence not constituting external bias, and for site-separable force fields each ΔE_i can be approximated simply and efficiently. In our self-guided variant of basin-hopping, particles are weighted by an approximate ΔE_i when randomly selected for an exchange, yielding a significant improvement for segregated multicomponent systems with modest particle-size mismatch.

Multicomponent clusters are of great contemporary interest for many applications, ranging from the use of alloyed metallic nanoparticles (*nanoalloys*)[1] in catalysis to the rational design of functional mesoscopic structures formed from colloidal pseudoatoms.[2] Theory and simulation have been productive in explaining and predicting the properties of both nanoalloys and colloidal clusters, with reliable structure prediction typically being a key objective. The global optimisation procedures required for model systems with chemical disorder are complicated by the combinatorial number of arrangements that may be possible for each underlying framework. In the present contribution we demonstrate a general and effective way to deal with this problem in a deterministic manner, which should expand the capability of theory and simulation to make efficient predictions and establish design principles for multicomponent systems.

Global optimisation for multicomponent systems is often split into two parts: (i) geometry optimisation via *displacement* moves and (ii) combinatorial optimisation of chemical ordering via *exchange* moves. A displacement move is an operation that alters the coordinates of at least one particle, without affecting the labels. Exchange moves do the converse: they interchange the labels of at least two unlike particles without affecting the coordinates; and they are often followed by a local relaxation of the geometry. The purpose of exchange moves is to find the most favourable permutational isomer (*homotop*[3]) for a given geometrical motif, and this is the issue addressed in the present contribution.

A common approach for homotop search is basin-hopping (BH)[4,5] with particle exchanges only.[6] With prior knowledge of the system, the randomly chosen exchange moves can be weighted in a way that biases the search and improves the efficiency.[7] Here we will show

that a deterministic algorithm based on Kernighan and Lin's (KL)[8] heuristic for partitioning graphs can be a more effective alternative. The heuristic is similar to the iterated local search (ILS) proposed by Lai *et al.*,[9] but with a number of subtle and important differences. We will also show how BH (with exchanges only) can be made *self-guided* by weighting individual particles in a way that does not constitute an external bias. For atom-separable potentials this weighting scheme can be approximated to achieve less reliable but computationally much more efficient results, providing a competitive alternative to the "tailored" exchange moves[7] that rely on chemical intuition. For the sake of brevity we shall restrict our discussion to binary systems, but the methodology can be applied to ternary and more complicated systems.

For a binary cluster with N_A particles of type A and $N_B = N - N_A$ of type B, each geometry might support up to $N!/(N_A! \times N_B!)$ distinct homotops. This number is often reduced by symmetry, but an exhaustive search for clusters with $N \gtrsim 35$ is still practically unfeasible. A Markovian random search without any guidance or bias is also not ideal, because it can potentially require even more than $N!/(N_A! \times N_B!)$ iterations due to multiple visits to the same configuration. On the other hand, it is clear that at most $\min(N_A, N_B)$ exchanges should ever be needed to reach one homotop from another. This observation implies that, if the most favourable homotop were known, it could be obtained from any other permutational isomer by exchanging no more than $\min(N_A, N_B)$ unlike pairs. Hence, a perfect search strategy would be the one that never requires more than $\min(N_A, N_B)$ swaps to find the best homotop. Also note that at every instance there are only $N_A \times N_B$ possible choices for a swap. Determining the "gain" (to be defined) associated with every possibility and using this information to guide the search becomes a feasible strategy. In the ideal case where swaps provide perfect guidance, the maximum number of possibilities that will need to be considered is

———————
[a)]Dmitri.Schebarchov@gmail.com
[b)]dw34@cam.ac.uk

$$N_A \times N_B \times \min(N_A, N_B) \quad \lesssim \quad N^3. \qquad (1)$$

Viewing a particle exchange as the basic step in combinatorial optimisation naturally leads to the notion of a *locally optimal* solution: it is one that cannot be improved by any single swap. Determining the number of locally optimal homotops would be a non-trivial task, but our intuition suggests that the number should be significantly smaller than $N!/(N_A! \times N_B!)$. It is also clear that the best (*globally optimal*) homotop must belong to this smaller set. Hence, there is potentially great utility in methods that can efficiently find a local optimum with respect to pair swaps, but not necessarily produce the global optimum. A powerful heuristic procedure due to Kernighan and Lin[8] (KL) achieves precisely this result.

The KL heuristic was intended as a rapid search for locally optimal partitions of an arbitrary graph, and the procedure involves *deterministic* swapping of individual nodes between subsets of given size. The optimality of a partition is measured by a specified "cost" function. Here we represent a multicomponent cluster as a generalised graph: the particles correspond to mobile nodes and the energy is the cost function, which does not have to be pairwise additive, and the graph edges need not be well-defined. The heuristic is well known to be effective in finding optimal solutions, and fast enough to be practical for solving large problems. The probability (p) that an optimal solution found by the KL procedure is globally optimal diminishes with system size (N), and the empirically-inferred scaling is approximately[8] $p(N) = 2^{-N/30}$. Assuming the same scaling applies to multicomponent molecular systems, for a 30-particle binary cluster at 50/50 composition we would expect the KL procedure to find the global minimum with $p \sim 0.5$. Here we shall not focus on the scaling with system size, but rather consider all possible compositions for one particular size to demonstrate the effectiveness of the KL procedure.

The KL heuristic starts from an arbitrary partition of particles into two indexed subsets $\mathcal{A} = \{a_i\}_{i=1}^{N_A}$ and $\mathcal{B} = \{b_j\}_{j=1}^{N_B}$, where each a_i and b_j is a label permanently fixed to a particle, and the corresponding potential energy is E. After initialising $E_O \leftarrow E$, $\mathcal{A}_O \leftarrow \mathcal{A}$, $\mathcal{B}_O \leftarrow \mathcal{B}$ and setting the termination condition C to $FALSE$, the procedure iterates according to the structure outlined in Fig. 1. The algorithm is guided by the quantities ΔE_{ij}, which we refer to as *swap gains*, each corresponding to the change in energy due to particles i and j trading places. The lists $\mathcal{A}'$ and $\mathcal{B}'$ are intended for tracking particles that are still in contention for a swap in the current *pass* (the **do** loop); and whenever two particles are exchanged, their labels are removed from $\mathcal{A}'$ and $\mathcal{B}'$. A pass will terminate when $\mathcal{A}'$ or $\mathcal{B}'$ is depleted. The procedure finishes when it fails to improve E_O in an entire pass, guaranteeing that the final $\mathcal{A}_O$ and $\mathcal{B}_O$ correspond to a locally optimal partition. Whenever an improvement is made during a pass, the best partition obtained during that pass is used to seed the next pass.

Remarkably, for all the weighted graphs considered by Kernighan and Lin (and all the model systems con-

```
while C is FALSE:
   Update C to TRUE.
   Set A' ← A and B' ← B.
   do min(N_A, N_B) times:
      Pick a'_I ∈ A' and b'_J ∈ B' with the lowest ΔE_IJ.
      Remove a'_I from A' and b'_J from B'.
      Move a_I to B and b_J to A.
      Update the total energy E.
      if E < E_O then:
         Update E_O ← E, A_O ← A and B_O ← B.
         Reset C to FALSE.
      end if.
   end do loop.
end while loop.
```

FIG. 1. Pseudocode for the KL procedure (see text).

sidered below), no more than four passes were ever required to converge on a locally optimal (or near optimal) partition.[8] This upper bound does not appear to be sensitive to system size, resulting in the total number of scanned exchanges scaling as $N^2 \log N$.[8] Note that it is the removal of particles from contention that improves the scaling from N^3 to $N^2 \log N$, because during each pass the number of swap possibilities is not fixed but rapidly diminishes from $N_A \times N_B$ to $|N_A - N_B|$. It is also important to note that swaps with positive as well as negative ΔE_{ij} can occur, which allows the system to escape from relatively shallow minima, thus improving the chances that the final partitioning is globally optimal. These two key features distinguish the KL heuristic from the "greedy" ILS of Lai *et al.*[9]

The computational bottleneck in the KL procedure is the selection of an exchange pair, which requires scanning all the possibilities to ensure the execution of the best swap every time. For weighted graphs with a pairwise-additive cost function the calculation of every ΔE_{ij} is relatively straightforward, and the total cost E can be updated very efficiently.[8] For an N-body system with a general energy function the bottleneck may be more severe. Most (semi-)empirical potentials for nanoalloys are not pairwise additive, in which case calculating each ΔE_{ij} may necessitate recomputing the total energy, and all of these computations will typically scale as N^2. Furthermore, exchanging two dissimilar particles is likely to result in undesirable local strain that should be relieved, and relaxing the geometry after every exchange is computationally expensive. To reduce this particular overhead in their ILS algorithm, Lai *et al.*[9] exploit a correlation between ΔE_{ij}^* *before* the relaxation and ΔE_{ij} *after* the relaxation, and they elect to rank all candidate pairs solely based on ΔE_{ij}^*. They then sequentially attempt exchanges according to this ranking, relax the corresponding structures and accept the first encountered improvement. This procedure reduces the number of local relaxations required to identify an exchange pair that warrants an improvement (not necessarily the best one). However, the performance of this strategy is expected to

deteriorate as the correlation between ΔE_{ij}^* and ΔE_{ij} gets worse (e.g. lattice mismatched systems).[10]

We now define the notion of *flip gain* (analogous to "cell gain" of Fiduccia and Mattheyses[11]), quantified by ΔE_i, which is the energy change due to a single particle i switching type (i.e. "flipping"). There are only N possibilities for a flip at any given instance: fewer than the number of unlike pairs, unless $\min(N_A, N_B) \leq 1$. The computational cost of selecting exchange candidates can be reduced by considering all the N flip gains instead of all the $N_A \times N_B$ swap gains. Fiduccia and Mattheyses[11] exploit this opportunity: they follow the same algorithm as in 1, but each choice of $a_I' \in \mathcal{A}'$ is determined by the lowest possible ΔE_I and each $b_J' \in \mathcal{B}'$ by the lowest ΔE_J. Although this strategy improves the scaling with system size, it can compromise the quality of the final solution. The potential error can be foreseen from the fact that particles I and J with minimal respective flip gains (ΔE_I and ΔE_J) will not necessarily yield the lowest possible swap gain (ΔE_{IJ}). In general, one cannot immediately infer the exact energy change resulting from a swap based on flip gains alone, especially if the potential is not pairwise additive and the two atoms are close to each other. Furthermore, relying on flip gains for guidance can result in desirable swaps being systematically missed, causing the algorithm to converge on a "good" solution that is not actually locally optimal. Note that one can select swap candidates based on flip gains calculated either in the same instance or in two successive instances. In the latter case, N_A and N_B are not strictly conserved, but each fluctuates by one mid-way through every swap, always reverting to the original values at the conclusion of each swap. Sequential updating guarantees the precise knowledge of the resulting change in total potential energy *before* a flip is executed: the change will always be equal to the corresponding flip gain. Of course, one still has to decide which particle type to flip first, and this choice may also affect the quality of the final solution.

If locating the globally optimal homotop is not the ultimate goal, but rather the aim is to find a homotop that is (in some sense) *almost* globally optimal, then relying on flip gains to select swap candidates would be the simplest strategy. By analogy with swap gains, for multicomponent clusters it is useful to distinguish between *approximate* flip gains (ΔE_i^*) *before* geometry relaxation and *exact* flip gains (ΔE_i) *after* the relaxation. Since most empirical potentials are site-separable, the total potential energy can be expressed as a sum over particles: $E = \sum_{i=1}^{N} E_i$; which admits efficient approximation of flip gains within a single function call. We chose to exploit this possibility by selecting swap candidates based solely on ΔE_i^*. The selection was performed sequentially with approximate flip gains updated mid-way through each swap. We also tried performing an additional mid-swap quench, but this procedure resulted in improvements that did not fully compensate for the cost of having twice as many geometry relaxations. Hence, in all the results that follow we essentially compare the following three schemes (now implemented in GMIN[12]):

KL The heuristic in 1 with a_I' and b_J' picked sequentially according to the minimum estimated flip gain, without mid-swap relaxation. The particle type with the lowest estimated flip gain was always picked first. When the search converged, it was restarted with another random initial condition.

BH* Self-guided basin-hopping with particles weighted by $\exp(-\Delta E_i^*/k_B T)$ in the random swap selection.

BH Standard basin-hopping with all particles weighted uniformly (i.e. random swaps with fair selection).

Although the general idea behind BH* is similar to "tailored" moves,[7] the intention is not to impose any external bias based on chemical intuition, but rather use the underlying energy landscape to make the search self-guided.

Our model system involves 22-particle binary Lennard-Jones (BLJ) clusters. This size was chosen because the global minimum structure[5] in the homogeneous case has relatively low symmetry (C_s), and all the homotops for every composition can be enumerated to confirm that the best homotop has indeed been found. The Lennard-Jones potential was chosen for its simple and generic form:

$$E = 4 \sum_{i<j} \epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{ij}} \right)^{6} \right], \qquad (2)$$

where $\alpha, \beta \in \{A, B\}$ specify the type of particles i and j, and r_{ij} is the distance between i and j. The parameters $\epsilon_{\alpha\beta}$ (the well depth) and $\sigma_{\alpha\beta}$ (particle radius) must also be specified, and the values considered in the present work never led to significant departure from the geometry of the homogeneous global minimum.[5] Starting from this structure, we first specify a composition and perform an extensive combinatorial search (10^6 swaps for BH and BH*, and 100 restarts for KL) using a particular method, with the geometry relaxed after every swap. Once the search has completed, we set the best encountered structure as a target, and then perform a hundred more searches using the same method, each terminating whenever the target is hit. All of the hundred searches were initialised with a different, randomly chosen homotop. The entire procedure was carried out for all compositions ($1 \leq N_A \leq 21$), and the search was performed using each of the three aforementioned schemes. The approximate flip gains were forecast in a single function call with transformed parameter matrices $\epsilon_{\alpha\beta}^*$ and $\sigma_{\alpha\beta}^*$:

$$\epsilon_{\alpha\beta} = \begin{bmatrix} \epsilon_{AA} & \epsilon_{AB} \\ \epsilon_{AB} & \epsilon_{BB} \end{bmatrix} \quad \mapsto \quad \epsilon_{\alpha\beta}^* = \begin{bmatrix} \epsilon_{AB} & \epsilon_{AA} \\ \epsilon_{BB} & \epsilon_{AB} \end{bmatrix}$$

and likewise for $\sigma_{\alpha\beta}^*$. The per-particle energies E_i^* accumulated during this call represent the energy each particle would have (*before* geometry relaxation) if it were (alone) to switch type. Subtracting the current per-particle energies E_i (computed with $\epsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta}$) will yield approximate flip gains: $\Delta E_i^* = E_i^* - E_i$.
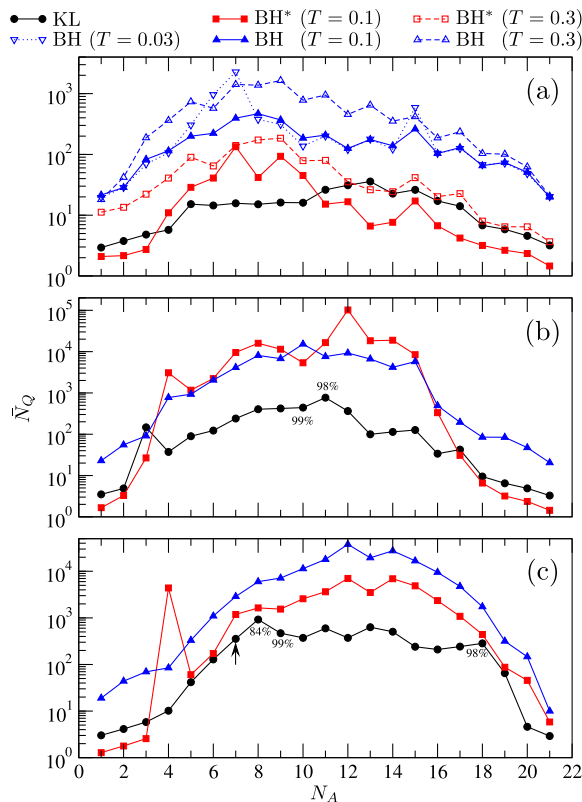
FIG. 2. Average number of swaps / quenches ($\bar{N}_Q$) required to hit the target homotop as a function of N_A — the number of type-A particles in the cluster. Three parameter sets were considered: (a) $\epsilon_{AB} = 1.2$, $\epsilon_{BB} = 1.5$; (b) $\epsilon_{AB} = 1.5$, $\epsilon_{BB} = 1.2$; (c) $\sigma_{AB} = 1.025$, $\sigma_{BB} = 1.05$; with the remaining $\epsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ in each case set to unity.

The hit statistics accumulated for three different parameter sets are shown in Fig. 2. The parametrisation considered in Fig. 2a is intended to represent a segregated system without lattice mismatch, in which case BH* and KL both require up to two orders of magnitude fewer swaps (and quenches) to find the global minimum. It is important to note that the performance of BH and BH* is sensitive to the temperature parameter in the accept/reject condition, which may require some tuning to achieve optimal performance, whereas KL has no adjustable parameters.

In Fig. 2b we again consider a lattice-matched system, but with well-depths set to encourage mixing. For this case KL outperforms both BH* and BH when $\min(N_A, N_B) \gtrsim 5$, i.e. the most challenging cases. Note that for $N_A = 10$ and $N_A = 11$ not all (99% and 98%) of the KL runs found the global minimum within the specified upper bound of one hundred restarts. Still, there is convincing evidence for KL outperforming the two variants of basin-hopping by two orders of magnitude.

The final parameter set considered in Fig. 2c represents a system of size-mismatched particles that are otherwise (chemically) equivalent. Note that the level of mismatch was chosen to ensure significant strain yet not alter the

geometrical motif of the global minimum.[13,14] Again, KL performs better than BH and BH*, especially for compositions near 50/50. However, for one data point in Fig. 2c (indicated by an arrow) KL failed to find the globally optimal solution in the initial search for a target. This result prompted us to repeat the run with more restarts, and we found that it located the global minimum only 6% of the time. In contrast, using exact flip gains leads to a hit-rate of 70%. This result exemplifies a potential pitfall: lattice mismatch can weaken the correlation between ΔE_i and ΔE_i^* to the point that the approximate flip gains can systematically misguide the search. The same caveat applies to approximate swap gains.[10] A detailed analysis of this issue and more benchmark calculations are in progress.

To summarise, we have outlined a deterministic scheme based on the Kernighan-Lin[8] heuristic and shown that it can outperform stochastic methods in combinatorial optimisation for binary clusters of 22 particles. We have also proposed a self-guided variant of basin-hopping, which improves conventional basin-hopping without relying on any external bias. We anticipate that both these methods be effective for combinatorial optimisation of a wide range of multicomponent clusters from the atomic to the mesoscopic length scale.

ACKNOWLEDGMENTS

[1] R. Ferrando, J. Jellinek, and R. L. Johnston, Chem. Rev. **108**, 845 (2008); *Metal nanoparticles and nanoalloys*, edited by R. L. Johnston and J. P. Wilcoxon (Elsevier, 2012); F. Calvo, A. Fortunelli, F. Negreiros, and D. J. Wales, J. Chem. Phys. **139**, 111102 (2013).

[2] G. Meng *et al.*, Science **327**, 560 (2010); D. Frenkel and D. J. Wales, Nature Materials **10**, 410 (2011); S. Hormoz and M. P. Brenner, Proc. Nat. Acad. Sci. USA **108**, 5193 (2011).

[3] J. Jellinek and E. Krissinel, in "emph "bibinfo booktitle Theory of Atomic and Molecular Clusters (Springer, 1999) pp. 277–308, ISBN 978-3-642-63570-0.

[4] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. USA **84**, 6611 (1987).

[5] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).

[6] L. O. Paz-Borbon *et al.*, Phys. Chem. Chem. Phys. **9**, 5202 (2007); J. Chem. Phys. **128**, 134517 (2008); G. G. Rondina and J. L. F. Da Silva, J. Chem. Inf. Model. **53**, 2282 (2013).

[7] D. Bochicchio and R. Ferrando, Nano Lett. **10**, 4211 (2010); Phys. Rev. B **87**, 165435 (2013).

[8] B. W. Kernighan and S. Lin, Bell Syst. Tech. J. **49**, 291 (1970).

[9] X. Lai, R. Xu, and W. Huang, J. Chem. Phys. **135**, 164109 (2011).

[10] M. Sicher, S. Mohr, and S. Goedecker, J. Chem. Phys. **134**, 044106 (2011).

[11] C. Fiduccia and R. M. Mattheyses, Proc. Nineteenth Design Automation Conf., 175(1982).

[12] D. J. Wales, "Gmin: A program for basin-hopping global optimisation, basin-sampling, and parallel tempering," .

[13] J. P. K. Doye and L. Meyer, Phys. Rev. Lett. **95**, 063401 (2005).

[14] D. J. Wales *et al.*, The Cambridge Cluster Database, URL http://www-wales.ch.cam.ac.uk/CCD.html(2001).