

# An improved sampling strategy for global energy minimization of multi-component systems



Richard Otis <sup>a,\*</sup>, Maria Emelianenko <sup>b</sup>, Zi-Kui Liu <sup>a</sup>

<sup>a</sup> Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA 16802, United States

<sup>b</sup> Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, United States

## ARTICLE INFO

### Article history:

Received 6 August 2016

Received in revised form 21 December 2016

Accepted 13 January 2017

Available online 2 February 2017

### Keywords:

CALPHAD

Thermodynamics

Energy minimization

Miscibility gap

## ABSTRACT

Effective initial sampling plays an important role in capturing key details about the energy surfaces of multi-component, multi-sublattice phases for the purposes of accurate convergence toward the global minimum energy configuration of a given system. It is shown that, when using the appropriate statistical distribution, both quasi-random and pseudo-random sampling methods compare well with the standard uniform grid-based technique. Moreover, the combination of random sampling with uniform grid points, while maintaining sampling performance for equilibrium calculations in the Al-Co-Cr system, significantly increases performance for a fictive 10-component system.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

Miscibility gap detection is a crucial feature in thermodynamic calculation software to accurately calculate the energy of phases containing regions of compositional instability and is commonly handled through global minimization (GM) of the Gibbs energy. The cause of miscibility gaps in non-ideal solutions is the presence of energetically unfavorable interactions between components that overwhelm the entropically-driven ideal mixing contribution to the Gibbs energy.

To fix notations, let  $G^i$  represent the molar Gibbs energy of a particular phase  $i$  within the system, with  $i = 1, \dots, K$ . Furthermore, assume that  $T$  is temperature,  $P$  is pressure,  $f^i$  is the fraction of phase  $i$  and  $y_{kj}^i$  is the site occupation fraction of component  $j$  in sublattice  $k$  of phase  $i$ . In vector form, let  $\vec{y} = \{y_{kj}^i\}_{j=1, \dots, J, k=1, \dots, M_i}^{i=1, \dots, K}$ , where  $K$  stands for the total number of phases,  $J$  for the total number of system components and  $M_i$  for the number of sublattices in phase  $i$ . Similarly, denote  $\vec{f} = \{f^i\}$  for all  $i = 1, \dots, K$ .

The isothermal-isobaric total molar Gibbs energy minimization problem for a closed system can be stated in the following way:

$$\min_{\vec{f}, \vec{y}} \left( G_m(T, P, \vec{f}, \vec{y}) = \sum_i f^i G^i(T, P, \vec{y}) \right) \quad (1)$$

$$\text{s.t. } c_n(T, P, \vec{f}, \vec{y}) = 0 \quad (2)$$

where  $c_n, n = 1, \dots, C$  represent all possible equality constraints including mass balance constraint. Sublattice site fractions are related to mole fractions by the relation

$$x_j^i = \frac{\sum_k b_k y_{kj}^i}{\sum_k b_k (1 - y_{k,va}^i)}, \quad (3)$$

where  $b_k$  is the number of sites on sublattice  $k$  and  $y_{k,va}^i$  is the fraction of vacancies on sublattice  $k$ .

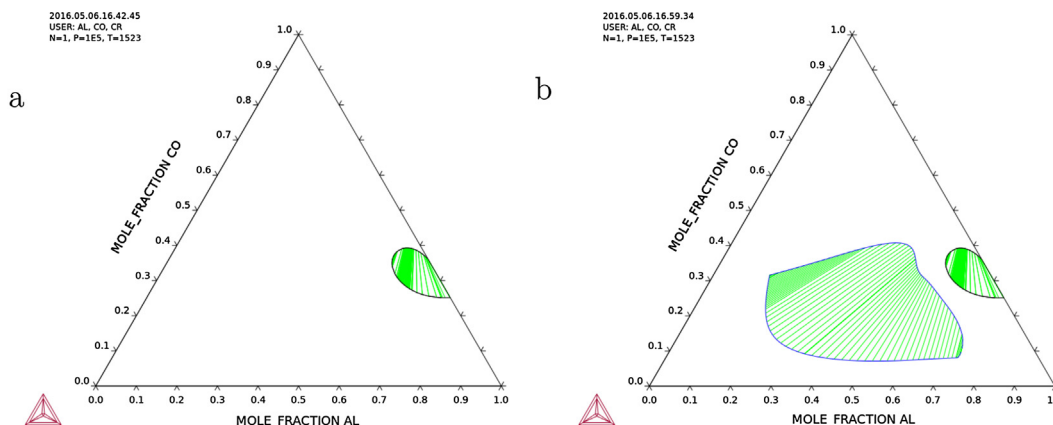
Site fraction balance constraints,  $\sum_k y_{kj}^i - 1 = 0$ , are always present for each sublattice. This problem is equivalent to solving the following unconstrained minimization problem for the Lagrangian  $\mathcal{L}$ :

$$\min_{\vec{f}, \vec{y}} \left( \mathcal{L}(T, P, \vec{f}, \vec{y}, \vec{\lambda}) = G_m(T, P, \vec{f}, \vec{y}) - \sum_{n=1}^C \lambda_n c_n(T, P, \vec{f}, \vec{y}) \right), \quad (4)$$

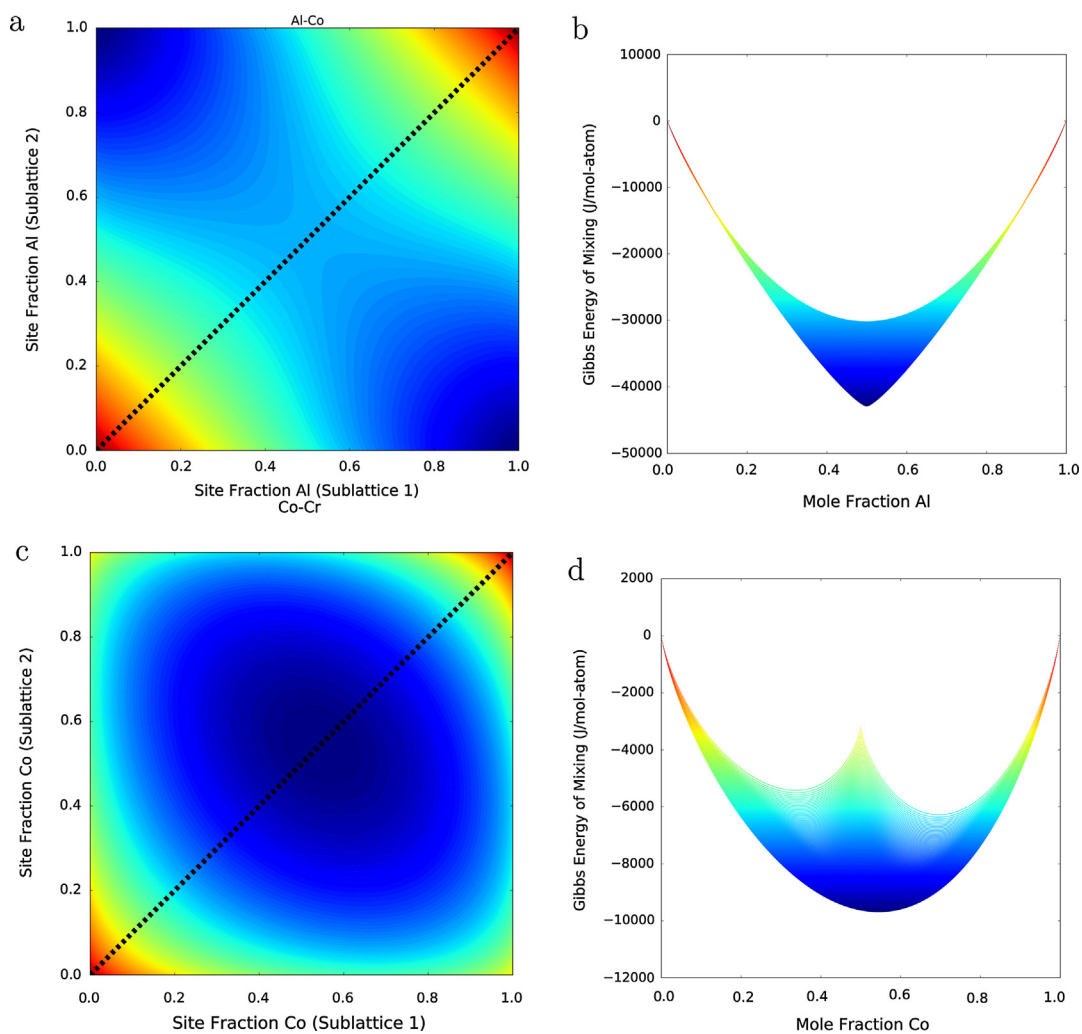
where the notation  $\vec{\lambda} = \{\lambda_n\}_{n=1, \dots, C}$  is used for the vector of Lagrange multipliers. The Lagrangian is closely related to the thermodynamic driving force function by the equality of the chemical potentials to the values of  $\lambda_n$  for the mass balance constraints. For further details,

\* Corresponding author.

E-mail address: [richard.otis@outlook.com](mailto:richard.otis@outlook.com) (R. Otis).



**Fig. 1.** Two Al-Co-Cr metastable phase diagrams, including only the bcc phase, are shown at 1523 K. Using the default settings in Thermo-Calc version 2015a, an incorrect diagram missing the ternary miscibility gap is produced (a). Increasing the global minimization point density or explicitly adding equilibrium points in the miscibility gap region produces the correct metastable diagram (b).



**Fig. 2.** The Al-Co (a, b) and Co-Cr (c, d) systems are shown with their internal energy surfaces (a, c) and those configurations mapped to the overall composition space (b, d). Multiple internal configurations map to the same overall composition, but only the low-energy configurations are relevant to equilibrium. For the Al-Co system, low-energy configurations are located at the Al:Co and Co:Al end-members, corresponding to the ordered bcc configuration of the phase near 50% Al. Due to crystallographic symmetry, Al:Co and Co:Al have the same energy, but this is not a requirement of sublattice models in the general case. The disordered bcc configuration, corresponding to the black dashed line across the diagonal, has higher energy in the composition region near 50% Al. Conversely, in the Co-Cr system, the ordered B2 end-members Co:Cr and Cr:Co are both much higher in energy than the disordered configuration of the same overall compositions, and that this holds true everywhere in that system at this temperature. This means B2 will not be observed at that temperature. The Al-Cr system (not shown) has a similar energy surface to the Co-Cr system.

interested readers are referred to the review of sublattices and the compound energy formalism found in [1].

From a computational perspective miscibility gaps pose a challenge, particularly for software employed in the application of the CALPHAD method [2,3], because they mean that the same phase may appear on multiple points on the equilibrium tie hyperplane but with different compositions. In these cases the minimization algorithm must increase the total degrees of freedom by creating multiple composition sets of the same phase, potentially up to the limit specified by the Gibbs phase rule. For multi-component systems the topology of the energy surfaces can become quite complex, since the high-dimensional tangent hyperplanes which define equilibrium can interact with the energy surface in more complicated ways. Moreover, when handling phases with sublattices, it is possible for points on the global energy surface (the composition) to be close together while being far apart in their site fraction coordinates (the constitution).

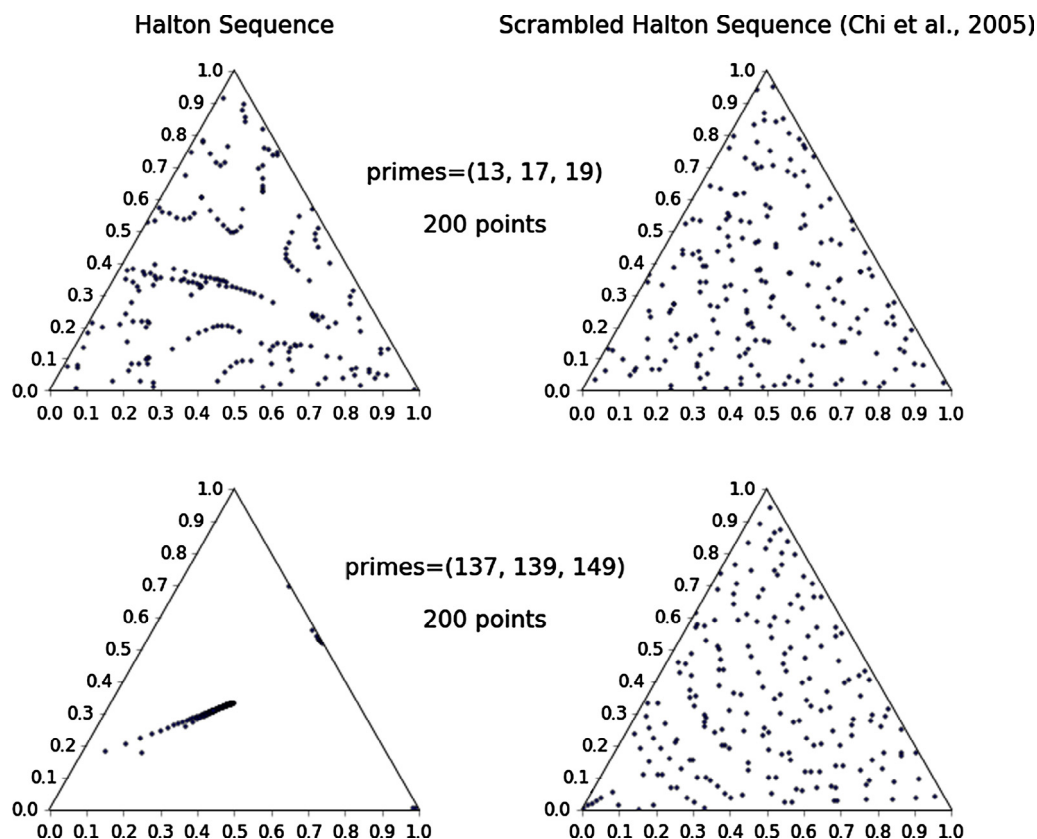
Few authors have discussed fully generalized GM schemes. Notable exceptions include the seminal works of Hillert [4] and Lukas [5], as well as more recently a discussion of the algorithms used in the OpenCalphad software [6]. In [7], a general constrained minimization methodology was proposed where all phases were duplicated to allow for miscibility gap detection. This strategy is easy to implement, but suffers an increase of complexity in high dimensions. The concept of using efficient sampling in this context was first introduced in low dimensions in [8]. Recently, there has been some work on tie hyperplane calculation for the multi-component case [9]. Other notable recent developments in the field of GM include the use of stochastic techniques such as particle swarm optimization [10,11], Cuckoo search [12], simulated annealing [13] and Tabu search [14].

Despite these advances, there has been limited discussion of methods for solving the multi-component case with multiple sublattices. A general method for this case is discussed in Section 3, without any attempts to compete with any of the aforementioned global minimization methodologies. The main focus of this work is in the development and testing of a multi-component, multi-sublattice sampling algorithm for generating an initial composition point set. This investigation is motivated by the fact that regardless of the choice of the GM solver, reliable convergence to the global energy minimum requires a good choice of starting points for the minimization procedure. The proposed strategy is relevant not only in the context of traditional optimization techniques such as Newton-Raphson method, but can help improve performance of the more sophisticated GM methods mentioned above. Development of such hybrid schemes is beyond the scope of the current investigations and will be explored in future work.

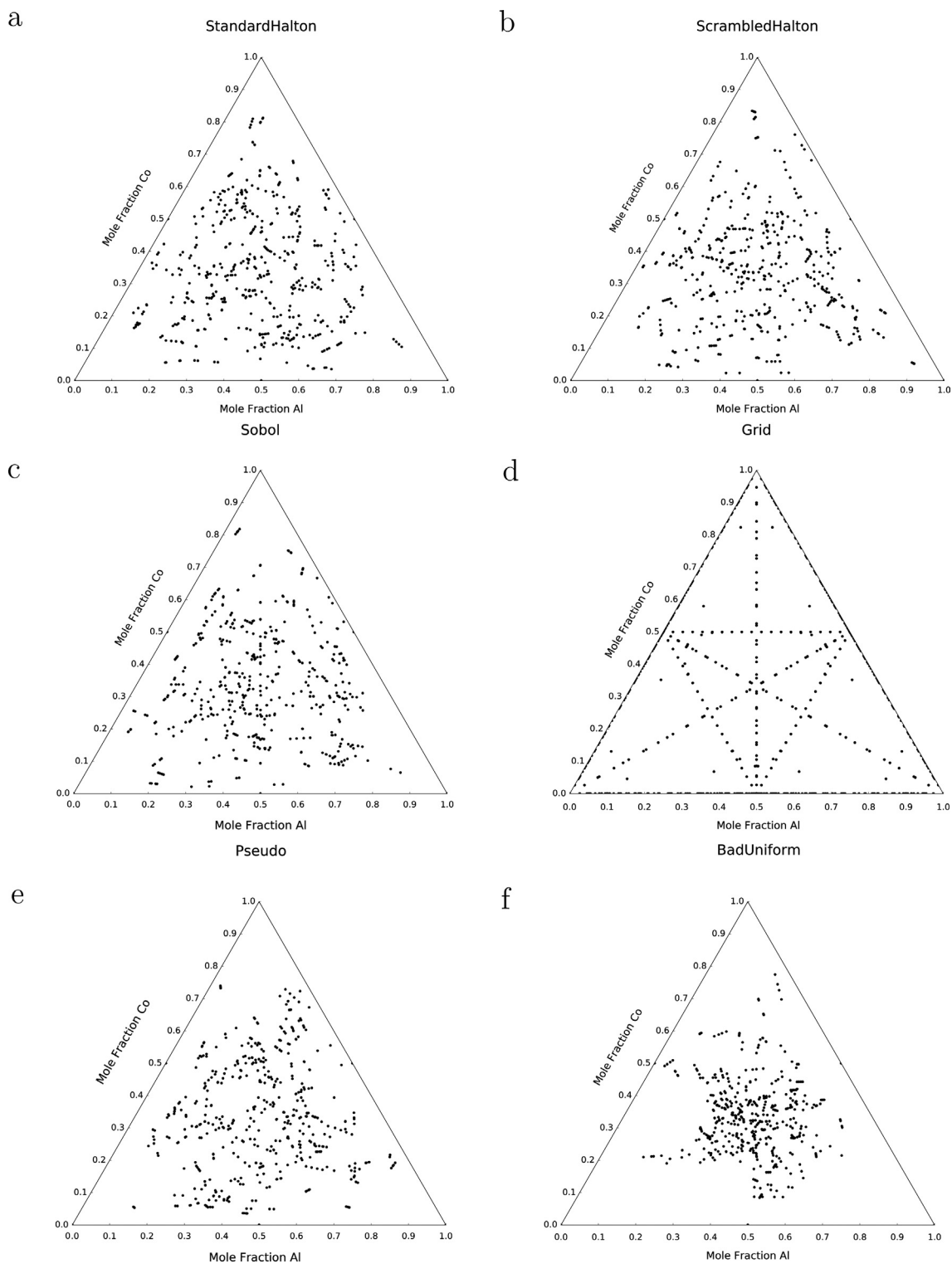
The paper is organized as follows. In Section 2 an example is provided of a practical system containing a miscibility gap where the deficiencies of current methodology become apparent. Section 3 is dedicated to fixing notations and clarifying the details of the GM method used in this work. Section 4 describes the novel quasi-random sampling strategy and benchmarks its performance against commonly used uniform sampling method. The results are summarized and conclusions are drawn in Section 5.

## 2. Motivation

A CALPHAD thermodynamic model for the Al-Co-Cr system was recently developed [15]. This system poses a challenge to CAL-



**Fig. 3.** The standard Halton sequence has problems with linear correlations between some dimensions. In some cases this causes significant undersampling of the space. The problem is relatively minor for the 6th, 7th and 8th prime bases (13, 17, 19) but becomes very apparent at high dimensions, as shown by the correlations between the 33rd, 34th and 35th prime bases (137, 139, 149) in the standard sequence. Using deterministic scrambling, these correlation effects in higher dimensions can be mitigated.



**Fig. 4.** Different sampling methods for global minimization are depicted for the two-sublattice bcc phase in the Al-Co-Cr system. Note that the BadUniform points (f) tend to cluster closer to the center of the triangle, while the Pseudo (e) and quasi-random methods (a–c) do a better job of covering the domain. The Grid method (d) has a distinct pattern that comes from projecting the higher-dimensional energy surface of the two-sublattice bcc phase onto the global composition triangle.

PHAD modeling software, despite only containing three components, in part due to a ternary miscibility gap in the bcc phase caused by complex atomic ordering. At 1523 K, for example, there is no problem plotting the stable diagram with all the phases, but a problem does emerge when considering the metastable bcc phase

diagram, shown in Fig. 1. Using the default settings in Thermo-Calc version 2015a, an incorrect diagram missing the ternary miscibility gap is produced. Increasing the global minimization point density or explicitly adding equilibrium points in the miscibility gap region produces the correct metastable diagram. It is unclear why this

occurs, and unfortunately the specific details of how phase diagram mapping works in Thermo-Calc are not publicly available, but the fact that these discrepancies exist at all motivates the present study of effective sampling.

It is first useful to understand how this bcc phase is modeled, according to [15]. The bcc sublattice model is (Al,Co,Cr,Va): (Al, Co,Cr,Va): (Va), with the third sublattice being empty and therefore irrelevant to the present discussion. Because this is effectively a two-sublattice phase, there are multiple internal configurations which will map to the same overall composition; this is the origin of the computational difficulty. This is challenging to depict for the full ternary system so, in Fig. 2, the internal energy surfaces (2a, c) are depicted for two of the constituent binary subsystems, as well as their mapping to the overall composition space (2b, d). For the Al-Co system (2a, b), low-energy configurations are observed at the Al:Co and Co:Al end-members, corresponding to the ordered bcc configuration (usually called “B2”) of the phase near 50% Al. Due to crystallographic symmetry, Al:Co and Co:Al have the same energy, but this is not a requirement of sublattice models in the general case. The disordered bcc configuration (“A2”), corresponding to the black dashed line across the diagonal, has higher energy in the composition region near 50% Al. Conversely, in the Co-Cr system (2b, d), the ordered B2 end-members Co:Cr and Cr:Co are both much higher in energy than the disordered configuration of the same overall compositions, and that this holds true everywhere in that system at 1523 K. This means B2 will not be observed at that temperature. The Al-Cr system, not shown, has a similar energy surface to the Co-Cr system.

Based on this analysis of the energy surfaces shown in Fig. 2 it is not surprising to observe the ternary miscibility gap, given that the stable configuration of the ternary bcc phase can involve a mixture of A2- and B2-type internal configurations. The next step is to develop a better understanding of how the sampling of the bcc phase’s internal configurations affects the algorithm’s convergence to the true equilibrium.

### 3. Numerical approach

Consider the multi-component, multi-phase case for fixed  $T$  and  $P$ , with the total mass of the system constrained to be  $N = 1$  mol. For a  $j$ -component system,  $j - 1$  additional constraints must be specified in order for the problem to be well-defined in accordance with the Gibbs phase rule. While many types of constraints can be specified, only specification of the mole fraction of a component,  $x_j$ , is explicitly considered here as a possible constraint. For this case, Algorithm 1 describes one possible approach to starting point

selection, mostly following Sundman [6,16], with a few insignificant differences in the implementation, as specified below.

#### Algorithm 1. GM starting point selection procedure

- 
- Setup. Initialize matrix  $M = I \in \mathbb{R}^{J \times J}$  (identity), chemical potentials  $\mu = (G_{max}, \dots, G_{max}) \in \mathbb{R}^{J \times 1}$  and specify the desired composition vector  $\vec{x}_0 = \{x_j\}_{j=1 \dots J}$ .
- Step 1. Sample a set of points  $\{y_{k,j}^{i,s}\}_{k=1 \dots M_i, j=1 \dots N_{i,k}}$ , where  $M_i$  is the number of sublattices in phase  $i$  and  $N_{i,k}$  is the number of sample points in  $k$ -th sublattice of phase  $i$ . Sampled points should include  $\{e_j\}_{j=1 \dots J}$  (corner points).
- Step 2. Compute the molar Gibbs energy values corresponding to all phases and sublattices:  $\{G^i(y_{k,j}^{i,s})\}_{i,j,k,s}$  forming a cumulative vector  $\vec{G} = \{G_q\}, q = 1, \dots, \sum_{i,k} N_{i,k}$ , with  $G_{max} = \max_q G_q$ . Compute site fractions by summing over sublattices  $x_j^{i,s} = \sum_k b_k y_{k,j}^{i,s} / (\sum_k b_k (1 - y_{k,Va}^{i,s}))$  (see (3)). Agglomerate sample points over all  $s = 1, \dots, N_{i,k}$  for a component  $j$  to form  $\{x_j^q\}$  representing all sample points  $x_j^q = \{x_j^{i,s}\}_{i=1 \dots K}$ . This set forms the computational grid used in Steps 3 and 4, which are aimed at determining the lower convex hull for this set. Initialize index set  $Q$  with  $J$  entries to contain indices  $q$  of the  $J$  corner points.
- Step 3. Compute the driving force for each sample point:  $\Delta G_q = G_q - \sum_{j=1}^J \mu_j x_j^q$ . Let  $q^* = \text{argmin}_q \Delta G_q$  (the most negative driving force), which picks the point  $G_{q^*}$  furthest below the line  $\sum_{j=1}^J \mu_j x_j^q$ .
- Step 4. Replace each row of  $M$  by  $x^{q^*}$ , one at a time, and solve  $M^T \vec{f} = \vec{x}_0$ . If all values  $f_j > 0$ , choose as the new  $M$  and update the replaced index with  $q^*$  in the set  $Q$ ; otherwise, try replacing the next row instead. Keep updating the index set  $Q$ . Recompute the chemical potentials by solving the linear system  $M \vec{\mu} = \vec{G}_Q$  to obtain new values  $\{\mu_j\}_{j=1 \dots J}$ .
- Step 5. If  $\|\mu^{new} - \mu^{old}\| < tol$ , terminate. Otherwise, go to Step 3.
- 

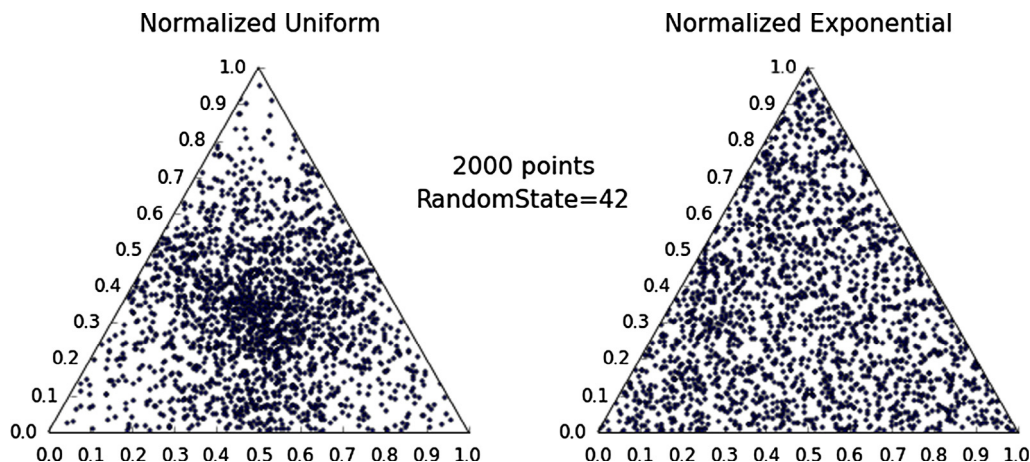


Fig. 5. Normalized uniform distributions produce poorer coverage of composition space versus the normalized exponential distribution. This is shown rigorously by Woronow [24]. The former distribution corresponds to the method called “BadUniform” in this work and, the latter, “Pseudo.”

Once a starting point has been selected, the local minimization problem can also be solved by constructing an augmented Hessian matrix including all phases' degrees of freedom and their constraints, as shown in (5):

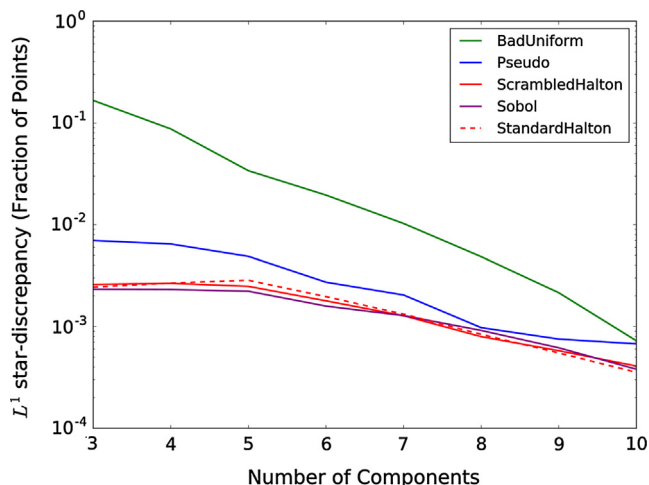
$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ \lambda_k \end{bmatrix} = \begin{bmatrix} -\nabla G_{m,k} + A_k^T \lambda_k \\ -c_k \end{bmatrix}, \quad (5)$$

where  $p$  is the Newton step direction for the iteration  $k$  and  $p_i$  is the step direction for the Lagrange multipliers  $\lambda_k$ .  $W_k$  is the Hessian matrix of  $L$  with respect to all degrees of freedom except  $\lambda_k$ ,  $A_k$  is the Jacobian matrix of the constraints,  $[\nabla c_1, \nabla c_2, \dots, \nabla c_m]^T$  and  $\nabla G_{m,k}$  is the gradient of the molar Gibbs energy of the system. This is sometimes called the “Newton-Lagrange” method and is detailed in standard texts, e.g., Section 18.1 of [17]. A similar approach for phase diagram calculation was reported by Lukas [5]. The key advantage of the latter approach is that you can optimize the degrees of freedom of all phases, including phase fractions and chemical potentials, simultaneously. The disadvantages are most apparent in large multi-component systems and systems with miscibility gaps; some procedure for adding and removing composition sets is required, in which case the system of equations must be rebuilt every time the stable set of phases changes. In the present implementation, this issue is addressed by eliminating phases of the same type when their sublattice site fractions become nearly equal during the iterative minimization; when the length of a tie-line approaches zero and the same phase is at both ends, the redundant phase is removed from consideration.

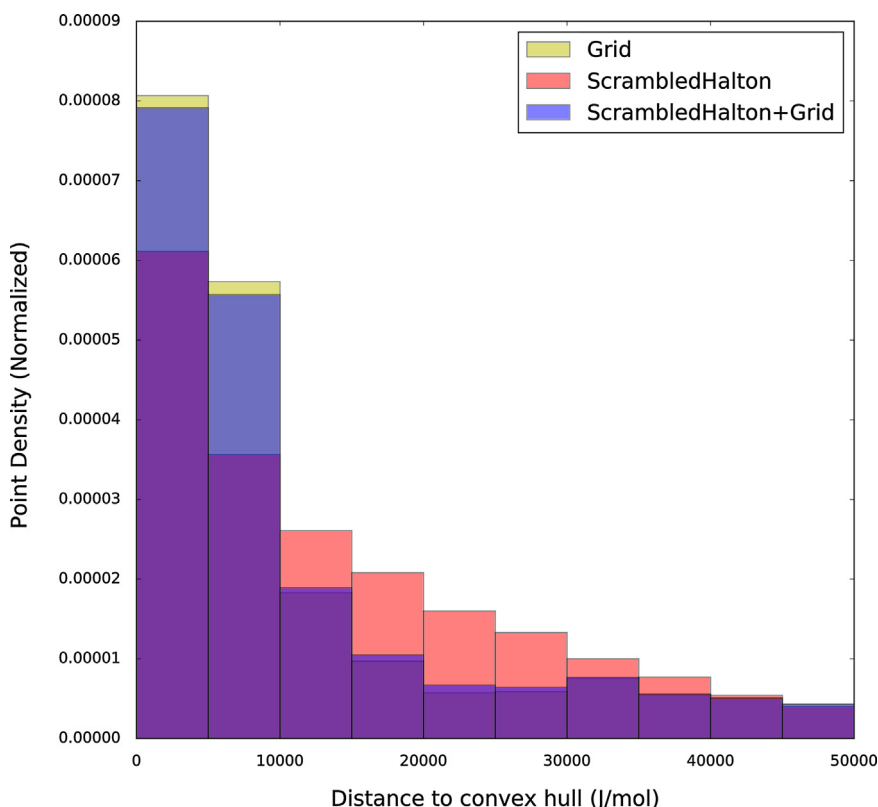
The main focus of this work is incorporating effective sampling strategies at Step 1 of the global minimization strategy described above. In what follows, several competing sampling strategies are compared.

#### 4. Sampling of composition space

Quasi-random sampling is a technique that attempts to augment the statistical advantages of pseudo-random sampling with



**Fig. 6.**  $L^1$  star-discrepancy calculation for 20000 points over the  $(N - 1)$ -simplex. A large number of points is required because the star-discrepancy measure is not sensitive enough to capture differences between the sampling methods in higher dimensions at low point density. The quasi-random sequences have consistently lower discrepancy than the pseudo-random samples. The “BadUniform” distribution shows significantly higher discrepancy due to the poor domain coverage of the method. The scrambled and standard Halton sequences do not show a significant difference up to ten components, since this is below the range where linear correlations appear.



**Fig. 7.** Grid points tend to lie closer to the equilibrium convex hull of the bcc phase than points sampled randomly. The fraction of sampled points within 10 kJ/mol of the equilibrium convex hull increases from 47% using scrambled Halton sampling to over 60% using either combined Halton and Grid sampling or pure Grid sampling.

the property of low “discrepancy,” meaning that the points will tend to evenly distribute throughout the domain. Unlike pseudo-random number generators, values from quasi-random sequences are explicitly designed to be biased, particularly toward spreading out to cover more of the domain. The quasi-random sequence can be constructed deterministically, and this is a desirable property from a reproducibility perspective because it ensures that, in principle, users with the same software version of the implementation will always converge to exactly the same solution.

The Halton sequence is an example of a quasi-random sequence of numbers generated over  $(0, 1)$  using a prime number as a base [18]. In low dimensions the  $N$ -dimensional Halton sequence is low-discrepancy, meaning that it is very similar or better than an equidistributed sequence in terms of domain coverage over  $(0, 1)^N$  for an equivalent number of points. Note that a sequence is referred to as equidistributed, or uniformly distributed, if the proportion of terms falling in a subinterval is proportional to the length of that interval. For larger  $N$ , the algorithm must be modified to prevent linear correlations between dimensions [19]; here the scrambling method recommended by Chi et al. [20] is used because of its effectiveness in reducing correlations, as shown in Fig. 3.

Similarly, the Sobol sequence [21] is another popular quasi-random sequence; here the Sobol sequence generator is seeded using the direction numbers recommended by Bratley and Fox [22], which are valid up to 40 dimensions, though, more recently, initializations up to 1111 dimensions have been developed [23]. Because both the Halton and Sobol sequences are extensible, mesh refinement routines can continually draw from the sequence by keeping track of the last computed member, without recomputing the entire sequence.

The quasi-random sequence alone does not meet the sampling requirement for composition space because it samples the  $N$ -hypercube, implying that all  $N$  dimensions are independent, but each sublattice’s site fractions must sum to unity in order to obey the site fraction balance constraint. Instead the  $(N - 1)$ -simplex must be sampled within each sublattice since the last coordinate is dependent. The most straightforward approach is to normalize each generated point so that each sublattice’s coordinates sum to unity, but this will tend to oversample the middle of the simplices, as shown in Fig. 5. Woronow [24] showed that uniformly distributed variates on the  $(N - 1)$ -simplex follow a  $N$ -dimensional symmetric Dirichlet distribution, which for the case of the full composition range simplifies to  $N$  samples from a standard exponential distribution divided by their sum. Exponentially distributed variates can be generated with the inverse transformation  $-\ln(U)$ , where  $U$  is a uniform random or quasi-random variate.

Because the Halton and Sobol sequences have the low discrepancy property, they can substitute for the uniform distribution; anywhere a uniform pseudo-random variate  $U$  would be used, it can be replaced with a value from a quasi-random sequence. This leads to the following general strategy for sampling, with specific methods subsequently listed.

#### 4.1. General simplex sampling strategy

1. Let  $N$  be equal to the total number of independent variables for a given phase.
2. Draw the first  $K$  points from an  $N$ -dimensional pseudo-random distribution or quasi-random sequence. Store the result in a  $K \times N$  matrix,  $M$ .
3. Compute  $-\ln(M)$  elementwise and store as the new  $M$ .
4. Group together columns of  $M$  corresponding to each sublattice. Compute the row-wise sum of each group and divide each element by its corresponding group’s sum.

For this work,  $K = 20(N - S)$ , where  $S$  is the number of sublattices.

#### 4.2. Sampling methods

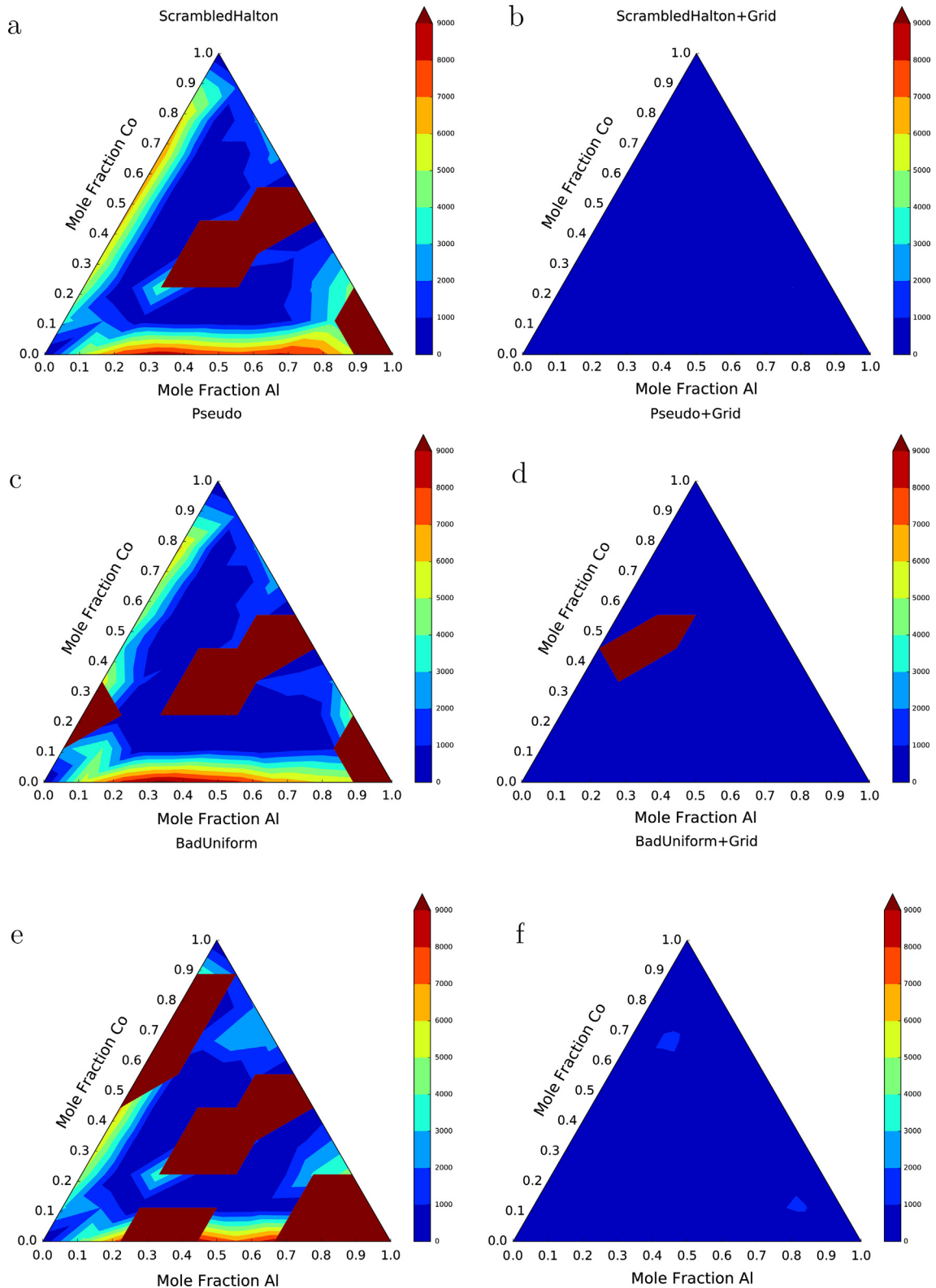
- **Pseudo:** Sample the uniform distribution.
- **BadUniform:** Same as Pseudo, but skip the negative-log transformation.
- **StandardHalton:** Choose the Halton sequence, using the first  $N$  primes.
- **ScrambledHalton:** Same as StandardHalton, but apply the scrambled method of Chi et al. [20]. This is often shortened to simply “Halton” when discussing the implementation.
- **Sobol:** Choose the Sobol sequence, seeded with the initialization numbers by Bratley and Fox [22].

In addition, these strategies are compared to a uniform grid-based sampling method referred to in this work as **Grid**, where evenly-spaced points are generated between all pairs of sublattice end-members. This is similar to the strategy employed by most CALPHAD software packages. (The exact strategy cannot be compared because the details are not published.) Fig. 4 visually depicts these methods for the bcc phase of the Al-Co-Cr system. (See Section 2 for details.) There are fewer sampled points at the corners of Al-Co-Cr because the mole fractions are not being uniformly sampled; rather, a higher-dimensional energy surface is being uniformly sampled and then projected back to mole fraction space. It is the clearest to see with the Grid sample subfigure that what are really being depicted are the features of a high-dimensional object, projected down to 2-D.

To quantitatively describe the deviation of a sampling method from an ideal lattice of points, which is not feasible to generate in higher dimensions, the discrepancy is computed. As noted earlier, the concept of discrepancy is that, for any arbitrary subregion of the domain, it is expected (or desired) that the fraction of sampled points contained within the subregion will be equal to the volume fraction of the subregion. The point fractions and volume fractions of all possible subregions are computed along with the absolute difference between the two, and then the maximum value is taken as the discrepancy of the point set. Searching all possible subregions is computationally difficult, especially in higher dimensions. For this comparison instead the “star-discrepancy” is computed, which restricts the subregions to be hypercube-shaped and anchored at the origin and each point. This generates  $K$  possible subregions and requires only  $O(K^2)$  computations for the point fractions. The fact that the space is actually a simplex poses no difficulty, since all the generated subregions using this approach will still be contained within the simplex.

Fig. 6 shows the result of an  $L^1$  star-discrepancy calculation for 20000 points using different sampling methods over the  $(N - 1)$ -simplex.  $L^1$  refers to the norm used to compute the difference between the volume and point fractions. A relatively high number of points are required, much higher than what is used in a typical equilibrium calculation, because the star-discrepancy measure is not sensitive enough to capture differences between the sampling methods in higher dimensions when the point density is low. The quasi-random sequences are shown to have consistently lower discrepancy than the pseudo-random samples. The “BadUniform” distribution shows significantly higher discrepancy due to the poor domain coverage of the method. The scrambled and standard Halton sequences do not show a significant difference up to ten components, since this is below the range where linear correlations appear.

For sampling in real systems, all sublattice end-members for each phase are also included by the algorithm to guarantee the convex hull of the points contains all of constitution space. An

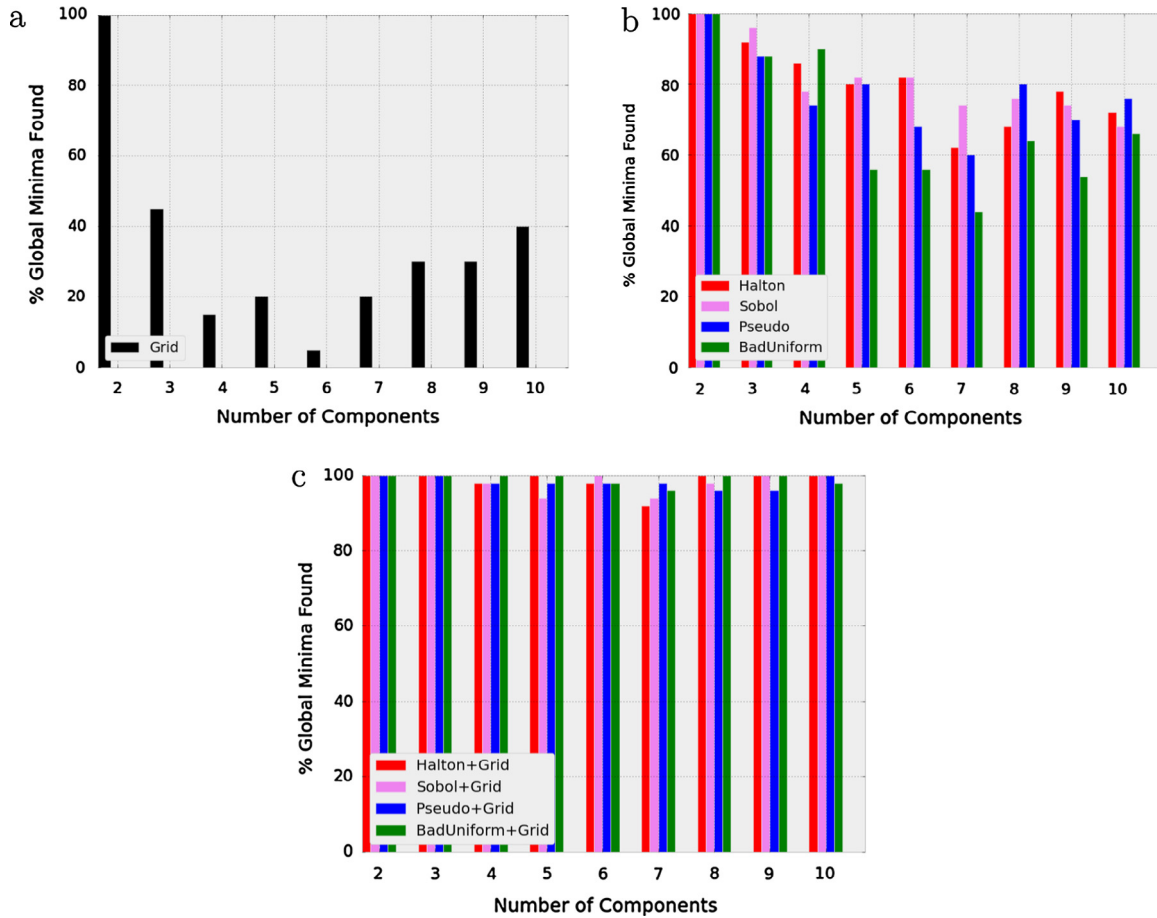


**Fig. 8.** The error in the Gibbs energy, computed by an equilibrium calculation at 1523 K in the Al-Co-Cr system, using the sampling schemes introduced in Section 4, is shown. The energy error is based on a comparison to a high-precision calculation, with the deep red color corresponding to a convergence failure. The BadUniform method (e) underperforms relative to the other methods. Pseudo (c) performs similarly to ScrambledHalton (a). The addition of points generated using the Grid method (b, d, f) reduces the error. The addition of points from the Grid method is beneficial to convergence because the equilibrium configurations in the bcc phase tend to lie along lines connecting pairs of end-members, the same lines which are sampled by the Grid method.

exception is the case of pure-vacancy end-members, since the molar Gibbs energy ( $G/N$ ) is not defined in the limit of pure vacuum ( $N \rightarrow 0$ ).

Another approach called rejection sampling [25], which involves sampling the  $N$ -hypercube and throwing out all points that fall outside the  $(N - 1)$ -simplex, is not considered in this work.





**Fig. 9.** The results of 50 randomly-generated trials of  $x^*$  are shown for numbers of components ranging from 2 to 10. The cases of pure Grid (a), pure random (b), and combined sampling (c) are considered. It is observed that, while the BadUniform method underperforms as expected, neither pure Grid nor any of the random sampling methods are able to find more than 80% of the global minima in the test systems when considering more than 6 components. The combined random + Grid strategy has significantly better results, with all methods finding at least 90% of the global minima, and all methods with the exception of BadUniform finding 100% of the 10-component global minima.

The problem with rejection sampling is that the fraction of rejected points quickly increases with  $N$ , making it very inefficient for multi-component systems.

#### 4.3. Performance of sampling methods

Fig. 8 shows the result of an equilibrium calculation at 1523 K in the Al-Co-Cr system using the sampling schemes introduced in Section 4. For this test the fcc and liquid phases are included along with bcc to make the test more realistic. The fully stable system also includes some intermetallics, which are neglected here. The color corresponds to the error in the total molar Gibbs energy when compared to a high-precision calculation, with the deep red color corresponding to a convergence failure. From the figure it is shown that the sampling methods perform roughly with the same trend as would be predicted by the discrepancies of their respective point sets, with the BadUniform method consistently underperforming relative to the other methods. The addition of points generated using the Grid method consistently reduces the error, and in many cases the reduction is to zero.

The addition of points from the Grid method is beneficial to convergence because the equilibrium configurations in the bcc phase tend to lie along lines connecting pairs of end-members, the same lines which are sampled by the Grid method. This is shown graphically in Fig. 7, where it can be seen that Grid

points tend to lie closer to the equilibrium convex hull of the bcc phase than points sampled randomly. The fraction of sampled points within 10 kJ/mol-atom of the equilibrium convex hull increases from 47% using scrambled Halton sampling to over 60% using either combined Halton and Grid sampling or pure Grid sampling. The combined sampling method (“random + Grid”) is simply the combination of points from a particular random sampling technique with points sampled by the Grid method.

It is not a surprising result to find that regular lattices of points are an effective sampling strategy. The reason that other approaches must be considered is that dense lattices are impractical to construct in higher dimensions due to the exponential increase in the number of required points.

Illustrating the value of random sampling with a real example is difficult due to the lack of publicly available multi-component CALPHAD databases with known, complex miscibility gap behavior, so instead a fictive system is constructed by adopting a global optimization test function [26], reproduced in Eq. (6) below. A similar test function was used to study global minimization performance in the CALPHAD software package Pandat [27].

$$f(x) = s \sum_{i=1}^n (x_i - x_i^*)^2 + \sum_{k=1}^{k_{\max}} a_k \sin^2 \left( f_k \sum_{i=1}^n (x_i - x_i^*)^k \right) \quad (6)$$

This function is an  $n$ -dimensional paraboloid, centered at  $x^*$ , with a second, strictly non-negative oscillating term. The second term can introduce curvature changes to the function's surface, making convergence to the global minimum more challenging. Because the second term is non-negative, it is clear that the global minimum value of this function must be zero, and that it takes on this value at  $x = x^*$ . (It is possible, for some choices of free parameters, for  $f(x)$  to have a value of zero at locations other than  $x^*$ .)  $n$  is the dimension of the space while  $s$ ,  $a_k$  and  $f_k$  are scaling parameters. The parameters are set  $k_{max} = 2$ ,  $f_k = 10$ ,  $a_k = 10^3$ ,  $s = 10^4$ .

Fig. 9 shows the results of 50 randomly-generated trials of  $x^*$  for numbers of components ranging from 2 to 10. The cases of pure Grid, pure random, and combined sampling are considered. The number of points used for each random sampling calculation is  $20(N - 1)$ , where  $N$  is the number of components, while the number of points for Grid sampling is  $20\binom{N}{2} = 20\sum_{i=1}^{N-1} i$ , but some Grid points will be repeated. It is observed that, while the BadUniform method underperforms as expected, neither pure Grid nor any of the random sampling methods are able to find more than 80% of the global minima in the test systems when considering more than 6 components. The combined random + Grid strategy has significantly better results, with all methods finding at least 90% of the global minima, and all methods with the exception of BadUniform finding 100% of the 10-component global minima.

## 5. Conclusion

Global energy minimization poses substantial theoretical and practical challenges to the CALPHAD modeling community. This work advances understanding of the role that effective sampling plays in capturing key details about the energy surfaces of phases, particularly multi-component, multi-sublattice phases, for the purposes of accurate convergence toward the global minimum energy configuration of the system. Both quasi-random and pseudo-random sampling methods compare well with the standard uniform grid-based technique. Moreover, the combination of random sampling with grid points, while maintaining sampling performance for a real-world multi-sublattice Al-Co-Cr system, also significantly increases performance for a fictive 10-component system. Further work is needed to understand how the performance advantage provided by a combined random + Grid sampling approach can be generalized to a greater variety of phase models found in CALPHAD modeling.

## Acknowledgements

This work was supported by a NASA Space Technology Research Fellowship grant NNX14AL43H, and by a National Science Foundation grant DMS-1056821.

## References

- [1] M. Hillert, The compound energy formalism, *J. Alloy. Compd.* 320 (2) (2001) 161–176, [http://dx.doi.org/10.1016/S0925-8388\(00\)01481-X](http://dx.doi.org/10.1016/S0925-8388(00)01481-X). <http://www.sciencedirect.com/science/article/pii/S092583880001481X>.
- [2] L. Kaufman, H. Bernstein, *Computer Calculation of Phase Diagrams*, Academic Press, 1970.
- [3] P. Spencer, A brief history of CALPHAD, *Calphad* 32 (1) (2008) 1–8, <http://dx.doi.org/10.1016/j.calphad.2007.10.001>. <http://dx.doi.org/10.1016/j.calphad.2007.10.001>.
- [4] M. Hillert, Some viewpoints on the use of a computer for calculating phase diagrams, *Phys. B+C* 103 (1) (1981) 31–40, [http://dx.doi.org/10.1016/0378-4363\(81\)91000-7](http://dx.doi.org/10.1016/0378-4363(81)91000-7). <http://linkinghub.elsevier.com/retrieve/pii/0378436381910007>.
- [5] H.L. Lukas, J. Weiss, E.-T. Henig, Strategies for the calculation of phase diagrams, *Calphad* 6 (3) (1982) 229–251, [http://dx.doi.org/10.1016/0364-5916\(82\)90004-9](http://dx.doi.org/10.1016/0364-5916(82)90004-9). <http://www.sciencedirect.com/science/article/pii/0364591682900049>.
- [6] B. Sundman, X.-G. Lu, H. Ohtani, The implementation of an algorithm to calculate thermodynamic equilibria for multi-component systems with non-ideal phases in a free software, *Comput. Mater. Sci.* 101 (2015) 127–137, <http://dx.doi.org/10.1016/j.commatsci.2015.01.029>. <http://www.science-direct.com/science/article/pii/S0927025615000361>.
- [7] J. Snider, I. Griva, X. Sun, M. Emelianenko, Set based framework for Gibbs energy minimization, *Calphad* 48 (2015) 18–26, <http://dx.doi.org/10.1016/j.calphad.2014.09.005>.
- [8] M. Emelianenko, Z.K. Liu, Q. Du, A new algorithm for the automation of phase diagram calculation, *Comput. Mater. Sci.* 35 (1) (2006) 61–74. <http://dx.doi.org/10.1016/j.commatsci.2005.03.004>.
- [9] N. Perevoshchikova, B. Appolaire, J. Teixeira, E. Aebly-Gautier, S. Denis, A convex hull algorithm for a grid minimization of Gibbs energy as initial step in equilibrium calculations in two-phase multicomponent alloys, *Comput. Mater. Sci.* 61 (2012) 54–66, <http://dx.doi.org/10.1016/j.commatsci.2012.03.050>. <http://linkinghub.elsevier.com/retrieve/pii/S0927025612001978>.
- [10] M.H.A. Piro, S. Simunovic, Global optimization algorithms to compute thermodynamic equilibria in large complex systems with performance considerations, *Comput. Mater. Sci.* 118 (2016) 87–96, <http://dx.doi.org/10.1016/j.commatsci.2016.02.043>. <http://www.sciencedirect.com/science/article/pii/S0927025616300829>.
- [11] H. Zhang, A. Bonilla-Petriciolet, G. Rangaiah, A review on global optimization methods for phase equilibrium modeling and calculations, *Open Thermodynam. J.* 5 (2011) 71–92.
- [12] A. Bonilla-Petriciolet, S. Fateen, G. Rangaiah, Constrained and unconstrained gibbs free energy minimization in reactive systems using genetic algorithm and differential evolution with tabu list, *Fluid Phase Equilib.* 340 (2013) 15–26.
- [13] A. Bonilla-Petriciolet, J.G. Segovia-Hernández, F. Castillo-Borja, U. Bravo-Sánchez, Thermodynamic calculations for chemical engineering using a simulated annealing optimization method, *Comput. Chem. Eng.* 24 (2007) 243–248.
- [14] Y. Teh, G. Rangaiah, Tabu search for global optimization of continuous functions with application to phase equilibrium calculations, *Comput. Chem. Eng.* 27 (2003) 1665–1679.
- [15] X.L. Liu, T. Gheno, B.B. Lindahl, G. Lindwall, B. Gleeson, Z.-K. Liu, First-principles calculations, experimental study, and thermodynamic modeling of the Al-Co-Cr system, *PLoS One* 10 (4) (2015) e0121386, <http://dx.doi.org/10.1371/journal.pone.0121386>. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0121386>.
- [16] B. Sundman, Private communication, 2014.
- [17] J. Nocedal, S. Wright, *Numerical Optimization*, second ed., Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006, <http://dx.doi.org/10.1007/978-0-387-40065-5>. <http://link.springer.com/10.1007/978-0-387-40065-5>.
- [18] J.H. Halton, Algorithm 247: radical-inverse quasi-random point sequence, 1964, doi:<http://dx.doi.org/10.1145/355588.365104>.
- [19] C. Schlier, On scrambled Halton sequences, *Appl. Numer. Math.* 58 (10) (2008) 1467–1478, <http://dx.doi.org/10.1016/j.apnum.2007.09.001>. <http://www.sciencedirect.com/science/article/pii/S0168927407001390>.
- [20] H. Chi, M. Mascagni, T. Warnock, On the optimal Halton sequence, *Math. Comput. Simul.* 70 (1) (2005) 9–21, <http://dx.doi.org/10.1016/j.matcom.2005.03.004>. <http://www.sciencedirect.com/science/article/pii/S037847540500087X>.
- [21] I. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals, *USSR Comput. Math. Math. Phys.* 7 (4) (1967) 86–112, [http://dx.doi.org/10.1016/0041-5553\(67\)90144-9](http://dx.doi.org/10.1016/0041-5553(67)90144-9). <http://www.sciencedirect.com/science/article/pii/0041555367901449>.
- [22] P. Bratley, B.L. Fox, ALGORITHM 659: implementing Sobol's quasirandom sequence generator, *ACM Trans. Math. Softw.* 14 (1) (1988) 88–100, <http://dx.doi.org/10.1145/42288.214372>. <http://dl.acm.org/citation.cfm?id=42288.214372>.
- [23] S. Joe, F.Y. Kuo, Remark on algorithm 659, *ACM Trans. Math. Softw.* 29 (1) (2003) 49–57, <http://dx.doi.org/10.1145/641876.641879>. <http://dl.acm.org/citation.cfm?id=641876.641879>.
- [24] A. Woronow, Generating random numbers on a simplex, *Comput. Geosci.* 19 (1) (1993) 81–88, [http://dx.doi.org/10.1016/0098-3004\(93\)90045-7](http://dx.doi.org/10.1016/0098-3004(93)90045-7). <http://www.sciencedirect.com/science/article/pii/0098300493900457>.
- [25] L. Devroye, General principles in random variate generation, in: *Non-Uniform Random Variate Gener.*, 1986, pp. 27–82, [http://dx.doi.org/10.1007/978-1-4613-8643-8\\_2](http://dx.doi.org/10.1007/978-1-4613-8643-8_2). [http://dx.doi.org/10.1007/978-1-4613-8643-8\\_2](http://dx.doi.org/10.1007/978-1-4613-8643-8_2).
- [26] J.D. Pintér, Global optimization: software, test problems, and applications, in: P.M. Pardalos, H.E. Romeijn (Eds.), *Handb. Glob. Optim., Nonconvex Optimization and Its Applications*, vol. 62, Springer, US, Boston, MA, 2002, pp. 515–569, <http://dx.doi.org/10.1007/978-1-4757-5362-2>. <http://link.springer.com/10.1007/978-1-4757-5362-2>.
- [27] W. Cao, S.-L. Chen, F. Zhang, K. Wu, Y. Yang, Y. Chang, R. Schmid-Fetzer, W. Oates, PANDAT software with PanEngine, PanOptimizer and PanPrecipitation for multi-component phase diagram calculation and materials property simulation, *Calphad* 33 (2) (2009) 328–342, <http://dx.doi.org/10.1016/j.calphad.2008.08.004>. <http://www.sciencedirect.com/science/article/pii/S0364591608000709>.