The Journal of Chemical Physics

RESEARCH ARTICLE | JULY 24 2024

Exact analytical solution of the Flory–Huggins model and extensions to multicomponent systems

J. Pedro de Souza ; Howard A. Stone

J. Chem. Phys. 161, 044902 (2024)
https://doi.org/10.1063/5.0215923

AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Exact analytical solution of the Flory–Huggins model and extensions to multicomponent systems

Cite as: J. Chem. Phys. 161, 044902 (2024); doi: 10.1063/5.0215923
Submitted: 26 April 2024 • Accepted: 17 June 2024 •
Published Online: 24 July 2024

J. Pedro de Souza[1,a] and Howard A. Stone[2,b]

AFFILIATIONS

[1] Omenn-Darling Bioengineering Institute, Princeton University, Princeton, New Jersey 08544, USA
[2] Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08544, USA

[a] Author to whom correspondence should be addressed: pdesouza@princeton.edu
[b] hastone@princeton.edu

ABSTRACT

The Flory–Huggins theory describes the phase separation of solutions containing polymers. Although it finds widespread application from polymer physics to materials science to biology, the concentrations that coexist in separate phases at equilibrium have not been determined analytically, and numerical techniques are required that restrict the theory's ease of application. In this work, we derive an implicit analytical solution to the Flory–Huggins theory of one polymer in a solvent by applying a procedure that we call the implicit substitution method. While the solutions are implicit and in the form of composite variables, they can be mapped explicitly to a phase diagram in composition space. We apply the same formalism to multicomponent polymeric systems, where we find analytical solutions for polydisperse mixtures of polymers of one type. Finally, while complete analytical solutions are not possible for arbitrary mixtures, we propose computationally efficient strategies to map out coexistence curves for systems with many components of different polymer types.

I. INTRODUCTION

Since its introduction in the early 1940s first by Huggins[1] and then by Flory,[2] the Flory–Huggins theory has become the most widespread model of phase separation of polymer solutions. Due to its simplicity in construction, the theory, with a minimal set of parameters, captures the main trends in the polymer partitioning between phases. The theory has been applied in numerous contexts, including in chemical processing,[3,4] materials science,[5–8] and also in the burgeoning field of biomolecular condensation.[9–11] While many corrections and modifications have been proposed to capture more detailed physical phenomena,[12–21] the original Flory–Huggins theory is the standard starting point for any analysis of materials that include polymeric components.

From its physics-based construction, the Flory–Huggins theory outputs a system of nonlinear constraints specifying thermodynamic equilibria that define the coexisting phase concentrations. While the equilibrium conditions can be solved numerically, up to this point, to our knowledge, no exact analytical solution has been found. Therefore, any evaluation of the theory or fitting of the theory to data requires numerical solutions, hampering the ease of use of the theory. In some regimes, analytical approximations can be applied,[22–26] but these approximations may involve complicated expressions and may not be uniformly valid, even for a single polymer–solvent combination.

The technical challenge of describing coexisting phases is more apparent for multicomponent mixtures,[27–29] even for mixtures of the same polymer type with different polymer lengths.[30] In order to construct a phase diagram using direct numerical solution, it is necessary to discretize the compositional space, and solve nonlinear systems of equations at each point. As the number of polymer types, M, increases, the discretized space increases exponentially, with a power proportional to M, which corresponds to a large number of discrete compositions, especially if fine resolution is desired. In many applications, such as in cellular biology, the number of polymer types can exceed tens of thousands.[31] Such a brute force

numerical approach may not be feasible to describe the breadth of coexisting concentrations over the full compositional space for $M > 4$, never mind for $M > 10^4$.

Here, we analytically derive an exact implicit solution to the Flory–Huggins theory of a one-polymer–one-solvent system using an implicit substitution method. The key idea is to solve for each constraint for the Flory–Huggins χ parameter, and then substitute the value of χ into each equation so as to eliminate it from the system. We explore the solution's asymptotic behavior, referring back to the original applications of the theory by Flory.[2] The general idea of our approach for this and other problems described in this paper is shown in Fig. 1.

Next, we apply the implicit substitution method to multicomponent mixtures where the relative magnitude of the pair interactions is known. There, we discover an exact analytical solution for polydisperse polymer samples of the same type, and distill the phase diagram of these mixtures into two independent variables that fully specify the coexistence curves. Using this strategy, we can analytically describe the composition (molecular weight distribution and total polymer volume fraction) of the coexisting phases.

For generalized mixtures, we use the implicit substitution method to simplify a nonlinear system of $M + 1$ equations down to one nonlinear "master" equation with one unknown—the solvent partitioning between the phases. While this reduction leads to great simplification of the equilibrium calculation, in the case of arbitrary mixtures, the master equation must be solved numerically. If desired, we can discretize the compositional space and construct a χ surface by solving just this one equation at each point—and we demonstrate this idea for a two-polymer–one-solvent mixture. To minimize the number of function evaluations, we propose a sampling method for the implicit function that obviates the need to discretize the entire compositional space. Choosing sampling points only requires finding the poles of the implicit χ function at a fixed distance from the global critical point, which we demonstrate, although we leave a full computational implementation of the compositional sampling for a large number of components to future work.

The implicit substitution method developed in this work is a powerful technique to solve or simplify the nonlinear systems of equations encountered in the thermodynamics of solutions. The only price to be paid is that the derived solutions are implicit in composite composition variables, themselves a function that is a combination of the coexisting phase concentrations. Nevertheless, these composite composition variables are experimentally accessible, and we hope that the technique will find applications in thermodynamic analyses in various contexts.

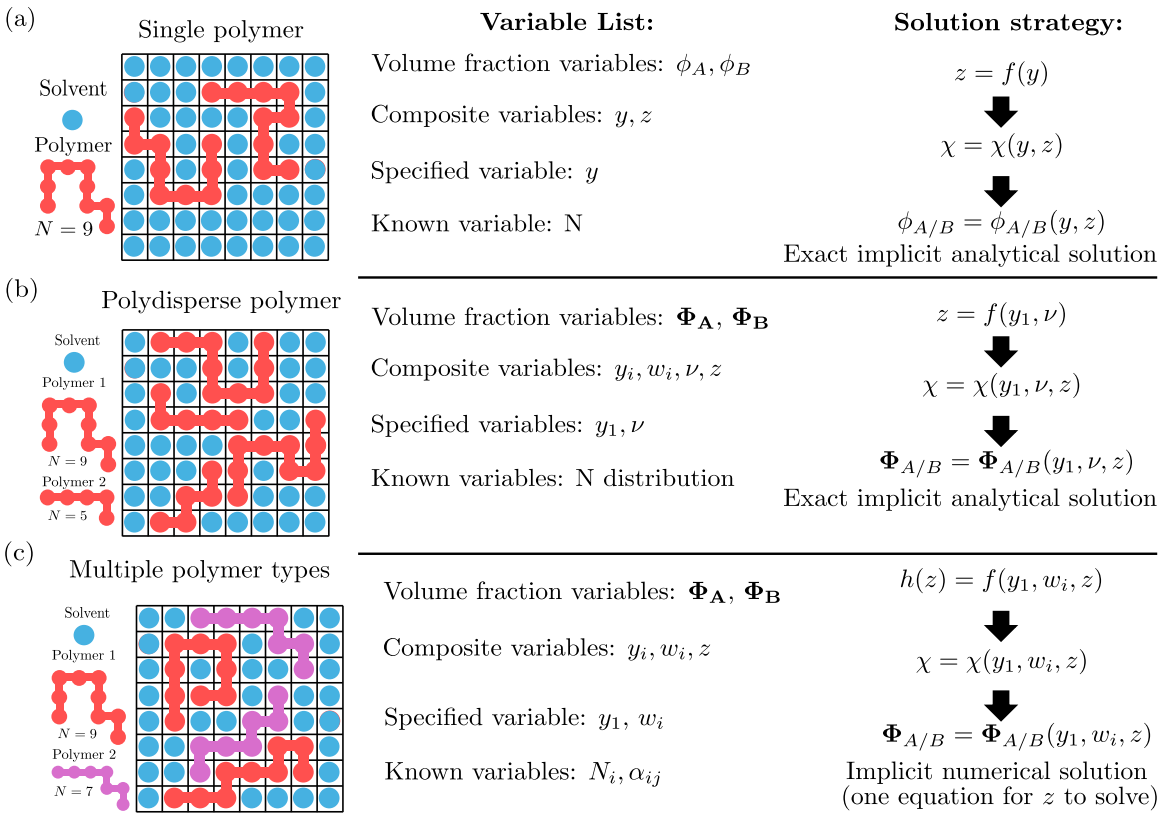

FIG. 1. Schematic of the systems under study, their corresponding variables, and the solution strategy employed in each case for (a) a single polymer–solvent solution, (b) a polydisperse polymer–solvent solution, and (c) a solution containing multiple polymer types in solvent.

II. THE FLORY–HUGGINS MODEL: ONE POLYMER

A. Theoretical background

The Flory–Huggins model may be derived using a free energy density where a polymer composed of N monomers occupies a lattice. The free energy density per lattice site, f, may be expressed in dimensionless terms, $\tilde{f}$, as

$$\tilde{f} = \frac{fv}{k_B T} = \frac{1}{N}\phi \ln(\phi) + (1-\phi)\ln(1-\phi) + \chi\phi(1-\phi). \tag{1}$$

Here, v is a lattice site volume, k_B is the Boltzmann constant, T is the absolute temperature, and ϕ is the polymer volume fraction. χ is the so-called Flory parameter, which captures the interaction between solvent and the polymer.[32]

From this free energy, we can define the dimensionless chemical potential of the polymer, $\mu = \partial\tilde{f}/\partial\phi$, up to an arbitrary constant,

$$\mu = \frac{1}{N}\ln(\phi) - \ln(1-\phi) - 2\chi\phi, \tag{2}$$

and the dimensionless osmotic pressure of the solution is $\Pi = -\tilde{f} + \phi\tilde{f}'$, so that

$$\Pi = \left(\frac{1}{N} - 1\right)\phi - \ln(1-\phi) - \chi\phi^2. \tag{3}$$

The model predicts two coexisting phases, a dense phase and a dilute phase of the polymer, above a critical χ value, $\chi > \chi_c$,

$$\chi_c = \frac{1}{2}\left(1 + \frac{1}{\sqrt{N}}\right)^2 \tag{4}$$

and beginning at a critical polymer volume fraction $\phi = \phi_c$,

$$\phi_c = \frac{1}{1+\sqrt{N}}. \tag{5}$$

It is found that as N increases, the phase diagram becomes asymmetric relative to the center line at $\phi = 1/2$.

Chemical equilibrium between two phases A and B at volume fractions ϕ_A and ϕ_B is specified by two constraints—equal chemical potential of the polymer and equal osmotic pressure of the solvent in the two phases,

$$\mu(\phi_A) = \mu(\phi_B), \tag{6a}$$

$$\Pi(\phi_A) = \Pi(\phi_B). \tag{6b}$$

These coupled equations can be rewritten explicitly as

$$\frac{1}{N}\ln\left(\frac{\phi_A}{\phi_B}\right) - \ln\left(\frac{1-\phi_A}{1-\phi_B}\right) - 2\chi(\phi_A - \phi_B) = 0, \tag{7a}$$

$$\left(\frac{1}{N} - 1\right)(\phi_A - \phi_B) - \ln\left(\frac{1-\phi_A}{1-\phi_B}\right) - \chi(\phi_A^2 - \phi_B^2) = 0. \tag{7b}$$

If χ is known, then we have two equations with two unknowns, ϕ_A and ϕ_B. To our knowledge, these coupled equations are impossible to deconvolve into two explicit functions describing ϕ_A and ϕ_B with known analytical techniques. However, in order to make analytical progress, we propose two simultaneous strategies. First, we search for implicit solutions, where we treat χ as an unknown, so as to gain an additional degree of freedom in the solution space; we can also eliminate χ from the system of equations quite easily. Second, instead of working with the variables ϕ_A and ϕ_B, we use two orthogonal composite variables that combine ϕ_A and ϕ_B, which we call y and z, which become apparent after eliminating χ. Thus, the problem is reduced to solving for one unknown variable, z, in terms of one known variable y. We call this solution method an implicit substitution method, since we use substitution to eliminate the function χ. Now that the strategy has been explained, we implement it in detail in the following.

B. Analytical solution

First, we solve Eqs. (7a) and (7b) for χ,

$$\chi = \frac{\frac{1}{N}\ln\left(\frac{\phi_A}{\phi_B}\right) - \ln\left(\frac{1-\phi_A}{1-\phi_B}\right)}{2(\phi_A - \phi_B)}, \tag{8a}$$

$$\chi = \frac{\left(\frac{1}{N} - 1\right)(\phi_A - \phi_B) - \ln\left(\frac{1-\phi_A}{1-\phi_B}\right)}{\phi_A^2 - \phi_B^2}. \tag{8b}$$

We equate these expressions to eliminate χ and multiply through by a factor $2(\phi_A^2 - \phi_B^2)$ to arrive at

$$\left[\frac{1}{N}\ln\left(\frac{\phi_A}{\phi_B}\right) - \ln\left(\frac{1-\phi_A}{1-\phi_B}\right)\right](\phi_A + \phi_B) = 2\left(\frac{1}{N} - 1\right)(\phi_A - \phi_B) - 2\ln\left(\frac{1-\phi_A}{1-\phi_B}\right). \tag{9}$$

Organizing the terms, we can write (9) as

$$\frac{(\phi_A + \phi_B)}{N}\ln\left(\frac{\phi_A}{\phi_B}\right) + (2 - \phi_A - \phi_B)\ln\left(\frac{1-\phi_A}{1-\phi_B}\right) = 2\left(\frac{1}{N} - 1\right)(\phi_A - \phi_B) \tag{10}$$

or

$$\frac{(\phi_A + \phi_B)}{N(\phi_A - \phi_B)}\ln\left(\frac{\phi_A}{\phi_B}\right) + \frac{2 - \phi_A - \phi_B}{\phi_A - \phi_B}\ln\left(\frac{1-\phi_A}{1-\phi_B}\right) = 2\left(\frac{1}{N} - 1\right). \tag{11}$$

If we define variables y and z as

$$y = \frac{\phi_A - \phi_B}{\phi_A + \phi_B}, \qquad z = \frac{\phi_A - \phi_B}{2 - \phi_A - \phi_B}, \tag{12}$$

Eq. (11) becomes

$$\frac{1}{Ny}\ln\left(\frac{1+y}{1-y}\right) - \frac{1}{z}\ln\left(\frac{1+z}{1-z}\right) = 2\left(\frac{1}{N} - 1\right). \tag{13}$$

The variables y and z are relative measures of the partitioning of polymer and solvent, respectively, between phases, and they are already commonly used as an order parameter that defines the phase composition differences[33] going back at least to Cahn and Hilliard.[34] The variable z is similar to y in that it contains the difference over the sum of solvent in each phase,

$$z = \frac{\phi_{sB} - \phi_{sA}}{\phi_{sA} + \phi_{sB}}, \tag{14}$$

where $\phi_s = 1 - \phi$ is the solvent volume fraction by incompressibility. It should be noted, however, that these mappings to y and z are not conformal. Nevertheless, since we have one degree of freedom, we will specify the variable y, and then solve the above-mentioned equation for z.

In this pursuit, we can first recognize the $\tanh^{-1}$ function, defined by the identity,

$$\tanh^{-1}(y) = \frac{1}{2}\ln\left(\frac{1+y}{1-y}\right) \tag{15}$$

to rewrite Eq. (13) as

$$\frac{\tanh^{-1}(y)}{Ny} - \frac{\tanh^{-1}(z)}{z} = \frac{1}{N} - 1. \tag{16}$$

If we define the function $h(x)$ as

$$h(x) = \frac{\tanh^{-1}(x)}{x}, \tag{17}$$

then Eq. (16) becomes

$$h(z) = 1 + (h(y) - 1)/N. \tag{18}$$

We call this function, $h()$, the "FH function" since it features prominently in our solution method for the Flory–Huggins model. Using the inverse FH function, $h^{-1}(x)$, which returns the positive branch, we can solve the equation for z in terms of y,

$$z = h^{-1}\left(\frac{1}{N}h(y) - \frac{1}{N} + 1\right). \tag{19}$$

Therefore, at this point, we have expressed the solution for the phase diagram exactly in terms of y and z. The function $h^{-1}(x)$ can be evaluated easily using predefined lookup tables (our strategy), although other strategies based on convergent series representations or asymptotic expressions of $h^{-1}(x)$ are possible. We leave these approaches to future practitioners if they are needed.

From the definition of y and z, we can find $\phi_A + \phi_B$,

$$\phi_A + \phi_B = \frac{2z}{y+z}. \tag{20}$$

Finally, we can return to Eq. (8b), expressing χ in terms of y,

$$\chi = \frac{\left(\frac{1}{N} - 1\right)(\phi_A + \phi_B)y - \ln\left(\frac{1-z}{1+z}\right)}{(\phi_A + \phi_B)^2 y}, \tag{21}$$

where we substitute in for z and $\phi_A + \phi_B$ in terms of y from Eqs. (19) and (20).

As mapped out in the schematic shown in Fig. 1(a), to construct the binodal curve, we specify a value of y between 0 and 1. Then, we find z by applying Eq. (19), then $\phi_A + \phi_B$ by applying Eq. (20), and then the corresponding χ value by applying Eq. (21). Note that we do not need to explore negative values of y since we arbitrarily assert that the A phase is enriched in the polymer relative to the B phase. If we wish to cast the results in terms of ϕ_A and ϕ_B, we can write

$$\begin{aligned} \phi_A &= \frac{z(1+y)}{z+y}, \\ \phi_B &= \frac{z(1-y)}{z+y}. \end{aligned} \tag{22}$$

If we enumerate multiple values of y between 0 and 1, we can construct the full phase diagram. Our solution is, therefore, an implicit function, $\chi(y)$, which we can map to $\chi(\phi)$.

The exact solutions to the Flory–Huggins theory are shown in Fig. 2 for $N = 10$, 50, 100, and 500. In Figs. 2(a) and 2(b), we show the functions $z(y)$ and $\chi(y)$ for each value of N, which we then map to the function $\chi(\phi)$ shown in Fig. 2(c). We confirm that the exact analytical solutions match numerical solutions of the coupled nonlinear equations.

Fitting of experimental data may be achieved directly by generating the full binodal curve and minimizing residuals from experimental data assuming some temperature dependence of χ. If both the dilute and concentrated phase concentrations are measured, then the experimental y can be computed directly and the values of $\chi(y)$ can be fitted *explicitly*. Using the variable y is also advantageous as it is independent of the assumed lattice volume size. To aid in the application of the theory, a supporting script written in Python is included to calculate the implicit curve $\chi(\phi)$.[35]

Surprisingly, along with the exact analytical solution, the approximation first used by Flory[2] is able to capture the asymptotic behavior of the phase distribution, especially for large N. It seems that this approximation has not received significant attention since most attention has been paid to analytical approximations in the region near the critical point and extensions from the critical point. For these reasons, we briefly explain the approximation method of Flory, as it shares some similarities to the implicit method we have applied and may be of practical use.

C. Extending Flory's approximation

Interestingly, in his 1942 article, Flory solved his theory in a rather ingenious way[2] using asymptotic arguments of the osmotic pressure difference between solutions. Recognizing that the osmotic pressure of the dilute phase would go to zero relative to the condensed phase, Flory solved the osmotic pressure equation for χ keeping only terms containing the condensed phase volume fraction. Specifically, he recognized that in the limit $N \gg 1$, and away from the critical point, it is natural to expect that $\phi_B \ll 1$, i.e., the solution is very dilute in polymer, and $\phi_A \approx 1$, i.e., the second phase is rich in polymer. In this limit, it follows that Eq. (7b) simplifies to

$$\left(\frac{1}{N} - 1\right)\phi_A - \ln(1 - \phi_A) - \chi\phi_A^2 = 0. \tag{23}$$

In other words, there is an implicit relation giving $\phi_A(\chi, N)$ in the form

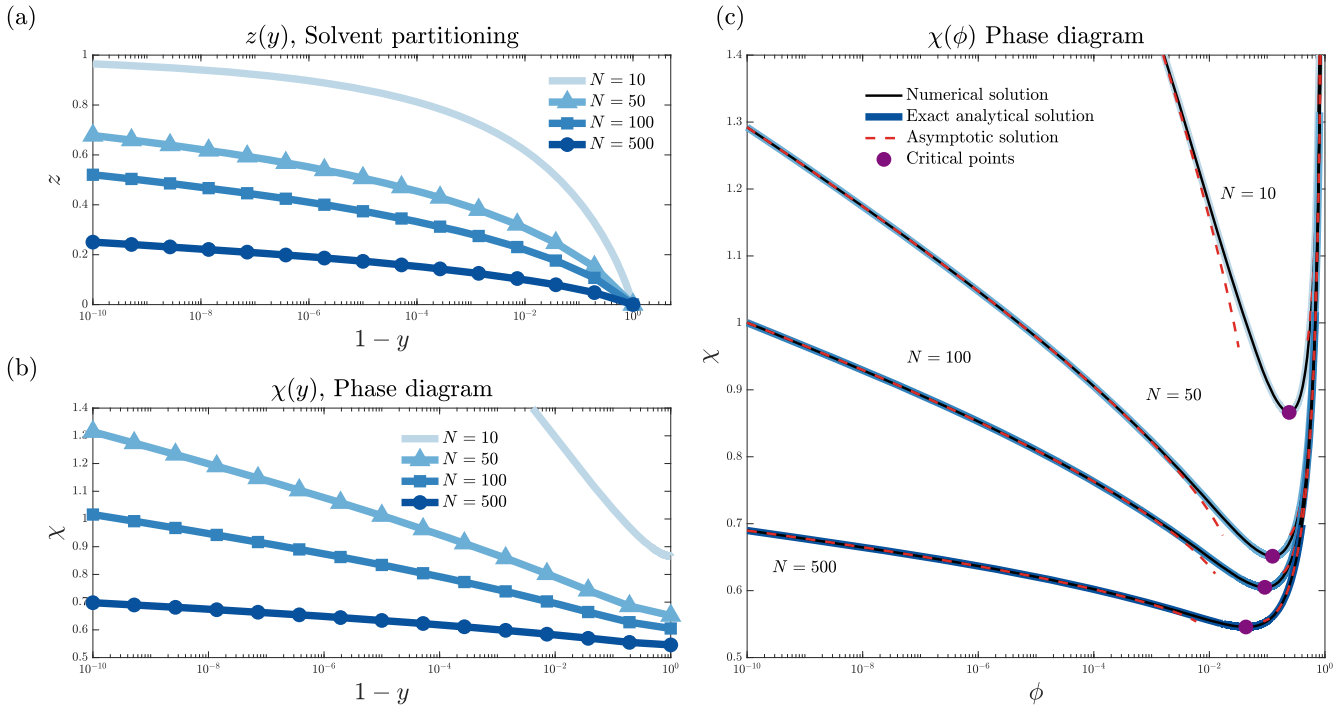

FIG. 2. Polymer–solvent solution. Phase diagram for a single polymer–solvent solution, comparing different degrees of polymerization ($N = 10, 50, 100, 500$). Panel (a) shows the analytical function $z(y)$ that describes the solvent partitioning as a function of the polymer partitioning between phases; panel (b) shows the analytical function $\chi(y)$; and panel (c) maps the coexistence curve to $\chi(\phi)$. The numerical solution overlaps perfectly with the exact analytical solution derived here. For large N, the asymptotic approximation first used by Flory[2] matches most of the curve far from the critical point. The purple circles mark the critical points at each value of N.

$$\chi = \frac{\left(\frac{1}{N} - 1\right)\phi_A - \ln\left(1 - \phi_A\right)}{\phi_A^2}. \tag{24}$$

Flory used this relationship in order to approximate the χ value that corresponds to a given dense phase concentration (ϕ_A) and then iteratively solved for the dilute phase concentration (ϕ_B) using the chemical potential constraint.

Interestingly, the starting approximation in Eq. (24) (with $\phi_A \gg \phi_B$) can be applied to the chemical potential Eq. (7a) directly to arrive at an implicit approximation for the dilute volume fraction, ϕ_B,

$$\frac{1}{N}\ln\phi_B = \frac{1}{N}\ln\phi_A - \ln\left(1 - \phi_A\right) - 2\chi\phi_A \tag{25}$$

or

$$\phi_B = \frac{\phi_A}{\left(1 - \phi_A\right)^N} e^{-2N\chi\phi_A}, \tag{26}$$

where χ for a specified ϕ_A is given in Eq. (24). This last result illustrates that typically ϕ_B is exponentially small.

As shown in Fig. 2(c), this asymptotic formula works incredibly well to describe the coexisting phases, especially for large N, with the exception of a small region near the critical point. Note that the approximation is computed for $2\phi_c < \phi_A < 1$, with ϕ_c defined by Eq. (5). Other approximations near the critical point are well documented, and they can also be extended to regions further from the critical point.[25,26]

Now that we have solved the single-polymer case, we can turn our attention to what we can learn about multicomponent mixtures.

III. MULTICOMPONENT MIXTURES

Most polymer solutions contain more than one polymeric species. Even for solutions containing a polymer of one type, the polymer sample often contains polymers of differing lengths (molecular weight), and each polymer of a specified length must be treated as a separate species.[30,36]

Here, we will introduce the extension of the Flory–Huggins model to multiple components. Then, the procedure for implementing the implicit substitution method for this system is expounded. Since the number of variables scales with the number of components, inevitably, we must introduce numerous composite variables along the way, which we define as needed. However, the solution strategy remains the same—working with variables z and y_i, using substitution of χ to arrive at one nonlinear equation with one unknown, and then, specifying only the allowed number of composite variables to arrive at a valid solution.

A. Theoretical background

For a multicomponent system with M polymeric species of volume fraction ϕ_i and size N_i in a solvent, the free energy density is[27]

$$\tilde{f} = \sum_{i=0}^{M} \frac{\phi_i}{N_i} \ln(\phi_i) + \frac{1}{2} \sum_{i=0}^{M} \sum_{j=0}^{M} \epsilon_{ij} \phi_i \phi_j, \tag{27}$$

where ϵ_{ij} encodes the symmetric interaction of species i with species j, and index 0 corresponds to the solvent, with $N_0 = 1$. Assuming incompressibility, $\sum_{i=0}^{M} \phi_i = 1$, we can substitute for the solvent volume fraction to arrive at

$$\tilde{f} = \sum_{i=1}^{i=M} \frac{\phi_i}{N_i} \ln(\phi_i) + \left(1 - \sum_{i=1}^{M} \phi_i\right) \ln\left(1 - \sum_{i=1}^{M} \phi_i\right) - \chi \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_{ij} \phi_i \phi_j, \tag{28}$$

where the sums are now over all M polymeric components (now not including the solvent—the standard notation from this point forward). The incompressibility constraint has been applied to eliminate the solvent density, and linear terms in ϕ_i are not included since they will not influence thermodynamic equilibrium. The coefficients α_{ij} indicate the shape of the symmetric effective interaction matrix, while the prefactor χ encodes the magnitude of interactions, such that $\chi\alpha_{ij} = -\frac{1}{2}(\epsilon_{ij} - \epsilon_{0i} - \epsilon_{0j} + \epsilon_{00})$. We assume that the matrix $\boldsymbol{\alpha}$ is known at a particular fixed temperature. The free energy reduces down to the form in Eq. (1) if there is only one polymeric species with $\alpha_{11} = 1$ and if the linear term in Eq. (1) is disregarded since it does not affect phase equilibrium calculations.

The chemical potential of each species, $\mu_i = \partial\tilde{f}/\partial\phi_i$, is up to an additive constant,

$$\mu_i = \frac{1}{N_i} \ln(\phi_i) - \ln\left(1 - \sum_j \phi_j\right) - \chi \sum_j 2\alpha_{ij}\phi_j. \tag{29}$$

Note that the chemical potential is defined relative to its reference state at fixed temperature and pressure.[37]

The osmotic pressure of the solution is computed as $\Pi = -\tilde{f} + \sum_i \phi_i \partial\tilde{f}/\partial\phi_i$, again up to an additive constant,

$$\Pi = \sum_i \left(\frac{1}{N_i} - 1\right)\phi_i - \ln\left(1 - \sum_i \phi_i\right) - \chi \sum_i \sum_j \alpha_{ij}\phi_i\phi_j. \tag{30}$$

We will assume that two phases A and B in equilibrium have composition $\boldsymbol{\Phi}_A$ and $\boldsymbol{\Phi}_B$, where $\boldsymbol{\Phi}$ indicates a vector with the components consisting of the corresponding chemical volume fractions. These phases satisfy the chemical equilibrium constraints of $\mu_{iA} = \mu_{iB}$ and $\Pi_A = \Pi_B$. Here, we will have $M + 1$ equations and $2M + 1$ unknowns ($\boldsymbol{\Phi}_A$, $\boldsymbol{\Phi}_B$, and χ). Therefore, we can only set M variables independently. Additional constraints may apply for the case of multiphase (>2) coexistence, but we will not consider this case on the outset. In fact, we will find that the two phase coexistence predictions can also be used as a way to find multiphase coexistence regions.

In order to make analytical progress, simplifying the set of nonlinear equilibrium constraints, we next cast the equations into a set of composite variables. A similar transformation was recently pursued for a different virial expansion model for multicomponent mixtures,[38] but these composite variables arise naturally from the Flory–Huggins construction. The following key composite variables are defined in Sec. III B: y_i, the partitioning of component i in Eq. (31); z, the partitioning of solvent in Eq. (32); β_i, the sum of the coexisting volume fractions of component i in Eqs. (32) and (44); γ_i, the differential of interactions of component i between phases in Eqs. (36) and (43); η_i, the differential of interactions of component i between phases relative to component 1 defined after Eq. (54); and finally w_i, the relative partitioning of component i relative to component 1 in Eq. (42). A reader who wishes to skip to the final working equations can skip to Eqs. (54) and (56), which together define the simplified master equation in one variable, z.

B. Casting into composite variables

Following the pattern from the one-component case with the implicit substitution method (Sec. II), we simplify the arguments of the ln functions by defining y_i as

$$y_i = \frac{\phi_{iA} - \phi_{iB}}{\phi_{iA} + \phi_{iB}}, \tag{31}$$

which are between −1 and 1; these variables represent the partitioning of component i between phases. Arbitrarily, we choose $0 \le y_1 < 1$, although all other y_i can take on negative or positive values.

The analogous variable for the solvent partitioning is denoted as z,

$$z = \frac{\sum_i (\phi_{iA} - \phi_{iB})}{2 - \sum_i (\phi_{iA} + \phi_{iB})}, \tag{32}$$

which again corresponds to the difference divided by the sum of volume fractions of solvent in respective phases, as in Eq. (14). If we define the set of variables β_i,

$$\beta_i = \phi_{iA} + \phi_{iB}, \tag{33}$$

as the sum of volume fraction of component i in coexisting phases, then we can relate z and y_i,

$$\sum_i y_i\beta_i = \left(2 - \sum_i \beta_i\right) z. \tag{34}$$

To start solving the system analytically, we first solve each constraint for χ. The M chemical potential Eq. (29) gives

$$\chi = \frac{\frac{1}{N_i} \ln\left(\frac{1+y_i}{1-y_i}\right) - \ln\left(\frac{1-z}{1+z}\right)}{\gamma_i} \quad \text{for every } i, \tag{35}$$

with the expression in the denominator, γ_i, defined as

$$\gamma_i = \sum_j 2\alpha_{ij}\beta_j y_j. \tag{36}$$

Therefore, γ_i is a measure of the differential of interactions of component i between phases.

From the osmotic pressure in Eq. (30), we get

$$\chi = \frac{\sum_i \left(\frac{1}{N_i} - 1\right)\beta_i y_i - \ln\left(\frac{1-z}{1+z}\right)}{\gamma_0} \tag{37}$$

with γ_0 defined as

$$\gamma_0 = \sum_i \sum_j \alpha_{ij}(\phi_{iA}\phi_{jA} - \phi_{iB}\phi_{jB}) = \sum_i \sum_j \alpha_{ij} y_j \beta_i \beta_j = \frac{1}{2}\sum_i \beta_i \gamma_i. \tag{38}$$

As in Sec. II, we can conveniently cast the equations for χ in terms of the $\tanh^{-1}$ functions,

$$\chi = \frac{\frac{2}{N_i}\tanh^{-1}(y_i) + 2\tanh^{-1}(z)}{\gamma_i} \quad \text{for every } i, \qquad \chi = \frac{2\tanh^{-1}(z) + \sum_i \left(\frac{1}{N_i} - 1\right)\beta_i y_i}{\gamma_0}. \tag{39}$$

For compact notation, we will define the variable x as

$$x = \tanh^{-1}(z), \tag{40}$$

so that $z = \tanh(x)$ and variables x_i as

$$x_i = \tanh^{-1}(y_i). \tag{41}$$

Before proceeding, we can simplify the expressions for γ_i [Eq. (36)] by defining a variable w_i, which is a measure of the relative partitioning of component i relative to the partitioning of component 1,

$$w_i = \frac{\beta_i y_i}{\beta_1 y_1} = \frac{\phi_{iA} - \phi_{iB}}{\phi_{1A} - \phi_{1B}}, \tag{42}$$

with $w_1 = 1$. The definition of γ_i in terms of the set w_i is, therefore,

$$\gamma_i = \beta_1 y_1 \sum_j 2\alpha_{ij} w_j, \qquad i \neq 0 \qquad \gamma_0 = \beta_1 y_1 \sum_i \sum_j \alpha_{ij} w_j \beta_i. \tag{43}$$

With the definition of w_i, by applying Eq. (34) we can solve for β_i in terms of the set of y_i, z, and w_i,

$$\beta_i = \frac{2 z w_i}{y_i \sum_j \left[w_j\left(1 + \frac{z}{y_j}\right)\right]}. \tag{44}$$

To summarize the mathematical manipulations thus far, we have recast the equations in terms of composite variables that simplify the arguments of the nonlinear logarithmic functions. Next, we work in terms of these composite variables, until we can write one master equation with one unknown by specifying a maximum of M independent composite variables.

C. Deriving the master equation

The next step, using our implicit substitution method, is to equate the equations for χ. For this purpose, we first equate χ from the first component to all the other components, to obtain $M - 1$ identities. Then, we equate a convenient linear combination of χ from all components to the χ specified by the osmotic pressure, in order to cancel the denominator γ_0. Proceeding, the combined chemical potential equations give

$$\frac{\frac{2}{N_i}x_i + 2x}{\gamma_i} = \frac{\frac{2}{N_1}x_1 + 2x}{\gamma_1}, \tag{45}$$

which is rearranged as

$$x_i = x\left(\frac{\gamma_i}{\gamma_1} - 1\right)N_i + \frac{\gamma_i}{\gamma_1}\frac{x_1 N_i}{N_1}. \tag{46}$$

Next, we apply the osmotic pressure constraint on χ, which we write as

$$\frac{2x + \sum_j \left(\frac{1}{N_j} - 1\right)\beta_j y_j}{\gamma_0} = \frac{\frac{2}{N_i}x_i + 2x}{\gamma_i}. \tag{47}$$

Multiplying this equation by $\beta_i \gamma_i$ and then summing over all components, we obtain

$$\frac{\sum_i \beta_i \gamma_i}{\gamma_0}\left(2x + \sum_j \left(\frac{1}{N_j} - 1\right)\beta_j y_j\right) = \sum_i \beta_i\left(\frac{2x_i}{N_i} + 2x\right). \tag{48}$$

Applying the definition of γ_0 so as to eliminate it, we find

$$\left(2x + \sum_j \left(\frac{1}{N_j} - 1\right)\beta_j y_j\right) = \sum_i \beta_i\left(\frac{x_i}{N_i} + x\right). \tag{49}$$

Then, collecting terms in x, we get

$$\left(2 - \sum_i \beta_i\right)x - \sum_i \beta_i y_i = \sum_i \left(\frac{\beta_i x_i}{N_i} - \frac{\beta_i y_i}{N_i}\right), \tag{50}$$

which can be written compactly as

$$\left(2 - \sum_i \beta_i\right)(x - z) = \sum_i \frac{\beta_i (x_i - y_i)}{N_i} \tag{51}$$

or as

$$\frac{\tanh^{-1}(z)}{z} = 1 + \frac{\sum_i \beta_i (\tanh^{-1}(y_i) - y_i)/N_i}{\sum_i \beta_i y_i}, \tag{52}$$

where we have reverted back to y_i and z from x_i and x. Furthermore, substituting in for β_i from Eq. (44) gives

$$\frac{\tanh^{-1}(z)}{z} = 1 + \frac{\sum_i w_i (\tanh^{-1}(y_i)/y_i - 1)/N_i}{\sum_i w_i}. \tag{53}$$

We can again write this expression in terms of the FH function, $h()$, defined by the following equation:

$$h(z) = 1 + \frac{\sum_i w_i (h(y_i) - 1)/N_i}{\sum_i w_i}. \tag{54}$$

Defining $\eta_i = \gamma_i/\gamma_1$, which is the differential measure of interactions between phases for component i relative to component 1, the chemical potential constraints in Eq. (45) give

$$x_i = N_i(\eta_i - 1)x + \eta_i N_i x_1/N_1 \tag{55}$$

or in terms of y_i,

$$y_i = \tanh\left(N_i(\eta_i - 1)\tanh^{-1}(z) + \eta_i N_i \tanh^{-1}(y_1)/N_1\right). \quad (56)$$

At this stage, we have reached our stated goal—to find one equation in one unknown by specifying at most M independent variables. With careful inspection, one may see that specifying y_1 and the set w_i for $i > 1$ (which constitutes M variables) is sufficient to achieve this objective. Stated more mechanically, by substituting Eq. (56) into Eq. (54), and specifying the M composite variables y_1 and w_i for $i > 1$, we have successfully transformed our system of $M + 1$ equations into one master equation, Eq. (54), in one unknown, z, the solvent partitioning between phases.

However, this equation does not have a clear inversion formula, since z is contained differently as an argument of multiple nonlinear functions, not just one nonlinear function as in the one polymer case. Therefore, in the most general case, the master equation must be solved numerically, but there are still significant advantages of this approach compared to the original numerical challenge, as we characterize in the following. However, it should be noted that there are clear strategies to approximate the master equation solution near and far from the critical point or when a particular pair interaction is dominant, as outlined in Appendixes A–D.

Next, instead of choosing w_i directly, we choose the set of $w_i y_1 = (\phi_{iA} - \phi_{iB})/(\phi_{1A} + \phi_{1B})$ for $i > 1$ and y_1, so that the specified variables have the same denominator of $\phi_{1A} + \phi_{1B}$, and so that they all go to zero at the global critical point. From w_i, one can determine the set η_i. Then, the master equation can be solved for z numerically. With z in hand, all the values of y_i and β_i can be calculated explicitly and then finally χ can be calculated using one of the original equilibrium constraints.

Even though a numerical solution of the master equation is still needed in general, a vast simplification has been achieved by all the algebraic manipulations entailed by the implicit substitution method—turning the system of $M + 1$ nonlinear equations into one nonlinear master Eq. (54) with one unknown. Again, we must work with the composite variables y_1 and w_i, but as shown in the schematic in Figs. 1(b) and 1(c), the solution can be directly transformed back to $\mathbf{\Phi}$ space with the relation,

$$\phi_{iA} = \frac{1}{2}\beta_i(1 + y_i), \quad (57a)$$

$$\phi_{iB} = \frac{1}{2}\beta_i(1 - y_i). \quad (57b)$$

In addition, we must also check that the solutions returned are actually physical, since admissible solutions to the master equation can return unphysical negative values of the volume fractions.

D. Special cases

In what follows, we discuss applications of the multicomponent solution method to three special cases. The first case is that of a single polymer with an arbitrary number of different lengths, i.e., the case of polydispersity. There, we find an exact analytical solution, and the master equation solution can be completely cast in terms of two independent variables without any numerical solution necessary. In the second case, we apply the solution method to a two-polymer–one-solvent system, to demonstrate how the method can be used to find the χ function over the full compositional space. Finally, in the third case, we suggest strategies to tackle the problem of large multicomponent systems of polymers using the implicit solution without discretizing the entire composition space.

1. Polydisperse polymer of one type

If a solution consists of only a solvent and only one polymer of one type, but that polymer is polydisperse with distinct values of N_i, then we can often assume the interactions are equal for each length of polymer in the mean field, i.e., $\alpha_{ij} = 1$, which means that $\eta_i = \gamma_i/\gamma_1 = 1$ for all i. In this case, the value of y_i becomes independent of z,

$$y_i = \tanh\left(\frac{N_i}{N_1}\tanh^{-1}(y_1)\right). \quad (58)$$

If this is the case, then we can again solve the equation for z explicitly with the inverse FH function h^{-1}, exactly as we did for the one-polymer case,

$$z = h^{-1}\left(1 + \frac{\sum_i w_i(h(y_i) - 1)/N_i}{\sum_i w_i}\right). \quad (59)$$

For arbitrary molecular weight distributions, the values of w_i and y_1 can be chosen independently to construct the full phase diagram. However, in many cases, the molecular weight distribution of the polymer being added to the solvent is known. If the average volume fraction of a particular molecular weight in both phases, $\bar{\phi}_i$, can be written in terms of v, the volume fraction of the phase A relative to the total volume,[30]

$$\bar{\phi}_i = v\phi_{iA} + (1 - v)\phi_{iB} = (\phi_{iA} - \phi_{iB})\left(v + \frac{1}{2y_i} - \frac{1}{2}\right), \quad (60)$$

then the ratio between the average polymer density $\bar{\phi}_i/\bar{\phi}_1$ is

$$\frac{\bar{\phi}_i}{\bar{\phi}_1} = \frac{\left(v + \frac{1}{2y_i} - \frac{1}{2}\right)}{\left(v + \frac{1}{2y_1} - \frac{1}{2}\right)} w_i. \quad (61)$$

Solving for w_i, we get

$$w_i = \frac{\bar{\phi}_i}{\bar{\phi}_1}\frac{\left(v + \frac{1}{2y_1} - \frac{1}{2}\right)}{\left(v + \frac{1}{2y_i} - \frac{1}{2}\right)}. \quad (62)$$

If y_1 and v are specified, then the set w_i is also fixed if the overall molecular weight distribution is known. Therefore, we can reduce the dimensionality of the phase diagram by only varying y_1 and v in the ranges $0 < y_1 < 1$ and $0 < v < 1$, as schematically shown in Fig. 1(b).

In Fig. 3, we show the phase diagram for two polymer samples of different overall molecular weight distributions, one that is uniformly distributed between $N = 10 - 100$ and one that is exponentially distributed between $N = 10 - 100$, where for both samples, the number of species is $M = 91$. These curves are analytically calculated in Fig. 3(a), which is much better than working with the 92 nonlinear chemical equilibrium equations that we started with. The curves are plotted as χ vs the overall polymer volume fraction, $\phi = \sum_i \phi_i$, at fixed condensed phase volume fraction, v. In Figs. 3(b)

and 3(c), the molecular weight distributions of the samples and the molecular weight distribution of the coexisting phases are plotted in the form of a probability density function. When the dense polymer phase volume fraction is small ($\nu \approx 0$), the binodal is pushed toward lower overall polymer density because the first appearing condensed phases preferentially include polymers of higher molecular weight, so they dominate the coexistence curve.

As the volume fraction of the condensed phase, ν, increases, the left branch of the coexistence curve (Fig. 3) shifts to the right since the condensed phase is forced to more closely match the overall molecular weight distribution as the condensed phase volume fraction increases. Although $\nu \approx 1$ may be difficult to reach, the minority dilute phase must enrich polymers of low molecular weight there since they are heavily crowded in the dense phase. Note that at $\nu = 0$ and $\nu = 1$, the molecular weight distribution of the dilute and dense phase, respectively, perfectly match the molecular weight of the overall polymer sample. Finally, comparing the two overall molecular weight distributions, the sample more enriched in the longer polymers has a left branch of the coexistence curve shifted to the left. In all the cases, the critical point appears to be dominated by the largest molecular weight polymer constituents, although it varies with the specific details of the molecular weight distribution and value of ν, as previously found in Ref. 30.

However, it should be noted that the volume fraction of the condensed phase, ν, may be difficult to estimate experimentally. In many cases, coexistence curves are mapped out upon the initiation of phase separation, which would occur at $\nu = 0$. Beyond this limit, for some practical cases where $\nu \neq 0$, the coexistence surface $\chi(\phi, \nu)$ could be more effectively (implicitly) mapped to the surface $\chi(\phi, \bar{\phi})$, where $\bar{\phi}$ measures the total volume fraction of polymer added to the solution, $\bar{\phi} = \nu\phi_A + (1-\nu)\phi_B$. Generating this surface requires a dense enumeration of possible values of ν, and therefore, we leave these extensions to future work.

In addition, similar approximations that reduce the compositional dimensionality can be made when one pair interaction is assumed to be dominant, as explained in detail in Appendixes C and D, to arrive at another set of analytical solutions.

2. A two-polymer-one-solvent mixture

While we were able to find an exact analytical solution for an arbitrarily polydisperse polymer, the two-polymer–one-solvent mixture has a master equation that can only be solved numerically. Here, the master equation becomes

$$h(z) = 1 + \frac{(h(y_1) - 1)/N_1 + w_2(h(y_2) - 1)/N_2}{1 + w_2}, \tag{63}$$

with y_2 fixed as

$$y_2 = \tanh\left(N_2(\eta_2 - 1)\tanh^{-1}(z) + \eta_2 N_2 \tanh^{-1}(y_1)/N_1\right) \tag{64}$$

and η_2 defined as

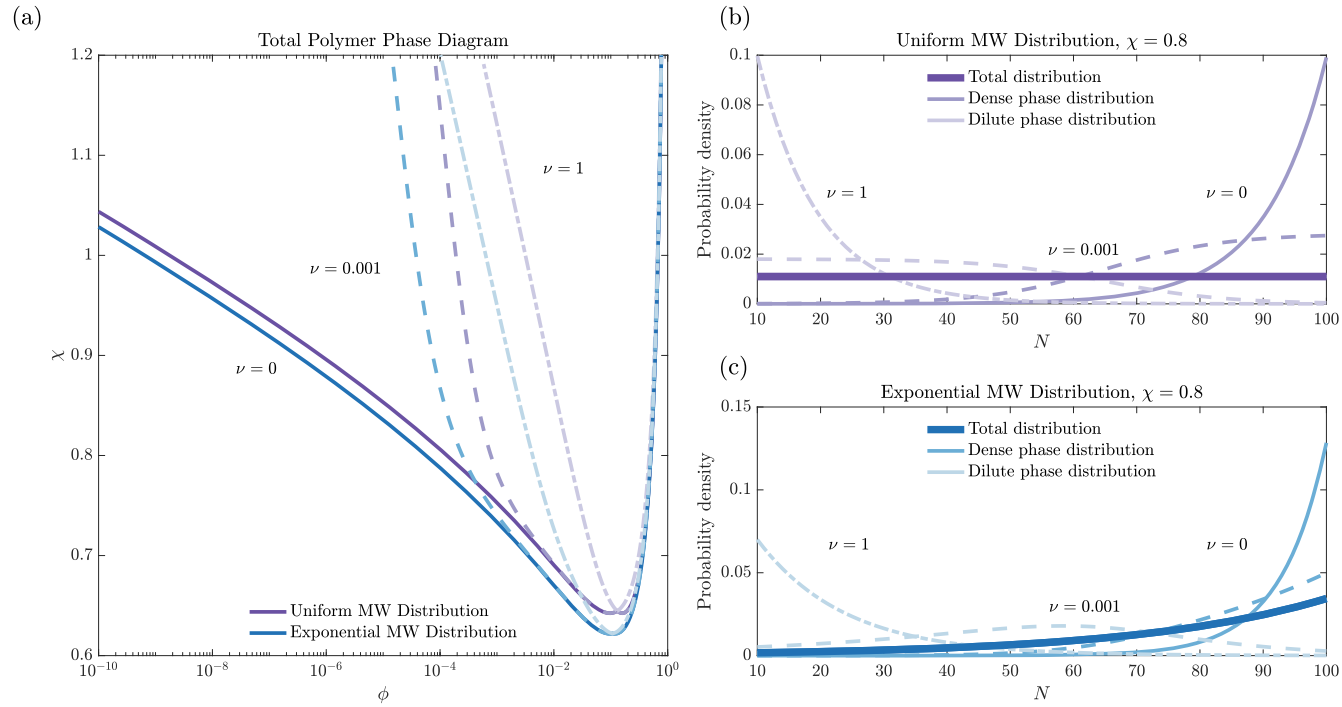

FIG. 3. Polydisperse polymer–solvent solution. Comparison of the phase diagram for two polydisperse samples composed of one polymer type. (a) Coexistence curves plotting χ vs total volume fraction of polymer $\phi = \sum_i \phi_i$. Each curve fixes the volume fraction of the condensed phase, ν. The binodal curves are calculated with a polymer with a uniform molecular weight distribution between $N = 10 - 100$ such that $M = 91$, and separately for a polymer with an exponential molecular weight distribution between $N = 10 - 100$, $\bar{\phi}_i/\bar{\phi}_1 = \exp\left((N_i - 10)/45\right)$. (b) and (c) The molecular weight distributions of the overall solution (thick lines), and the molecular weight distributions in the dense and dilute phases at $\chi = 0.8$. Note that for $\nu = 0$ and $\nu = 1$, the dilute and dense phases, respectively, perfectly overlap with the overall molecular weight distribution.

$$\eta_2 = \frac{\alpha_{12} + \alpha_{22} w_2}{\alpha_{11} + \alpha_{12} w_2}. \tag{65}$$

To construct a phase diagram, we specify y_1 and $y_1 w_2$, then solve for η_2, and then for z solving the master equation numerically. From there, we can specify y_2, χ, and then solve explicitly for the polymer densities in each phase. To discretize the phase diagram, we vary $0 < y_1 < 1$ and $-1 < \tanh(y_1 w_2) < 1$. It should be noted that it is more regularly behaved to plot the phase diagram in terms of y_1 and the product $y_1 w_2$, since they share the same denominator in terms of their $\mathbf{\Phi}$ dependence.

In Fig. 4, we present the solution for a two-polymer–one-solvent system for two different shapes of the interaction matrix: panels (a)–(c) show a diagonally dominant interaction matrix and panels (d)–(f) show an off diagonal dominant matrix. So as to not push the coexisting densities to extreme values near 0 or 1, we choose $N_1 = 4$ and $N_2 = 3$ as a starting point. Note that the same method may be applied for larger values of N_i, but care must be taken in order to include values of y_1 near 1. The χ surface is calculated over the full compositional landscape. Interestingly, all differentiable regions emanate from the central critical point at $y_1 = 0$ and $w_2 y_1 = 0$, but are separated by singular discontinuous curves in the phase diagram. Not all of these regions or parts of these regions return physical, positive densities, as shown in Figs. 4(b) and 4(e). Nevertheless, by transforming specified contours with admissible solutions to $\mathbf{\Phi}$ space, we can map out the phase coexistence curves in terms of ϕ_1 and ϕ_2, as shown in Figs. 4(c) and 4(f).

There are a few notable features of the full χ surfaces, particularly in Fig. 4(b). First, some contours do not connect to the global critical point, and these coexistence curves at fixed χ, therefore, lack a critical point within the physical values of $\mathbf{\Phi}$ at this fixed χ, for example, $\chi = 1.8$. Furthermore, there are multiple contours that have the same value that are separated by nonphysical and nondifferentiable regions of the χ surface, for example, $\chi = 3.4$. These separate contours indicate phase coexistence beyond two phases. However, these particular contours are localized in smaller regions in $\mathbf{\Phi}$ space, and they are not as illustrative as the ones chosen for Fig. 4(c). Nevertheless, using this method, which has been constructed for the coexistence of *two phases*, we can construct phase diagrams in which more than two phases are present where tie lines cross at fixed χ. We will investigate these multiphase predictions particularly in a future work.

Clearly for $M = 2$, if the value of the desired χ contour is known, mapping out the full χ landscape may be more computationally expensive than solving the equations in ϕ space for the fixed, desired value of χ. So, the fully discretized χ surface implemented

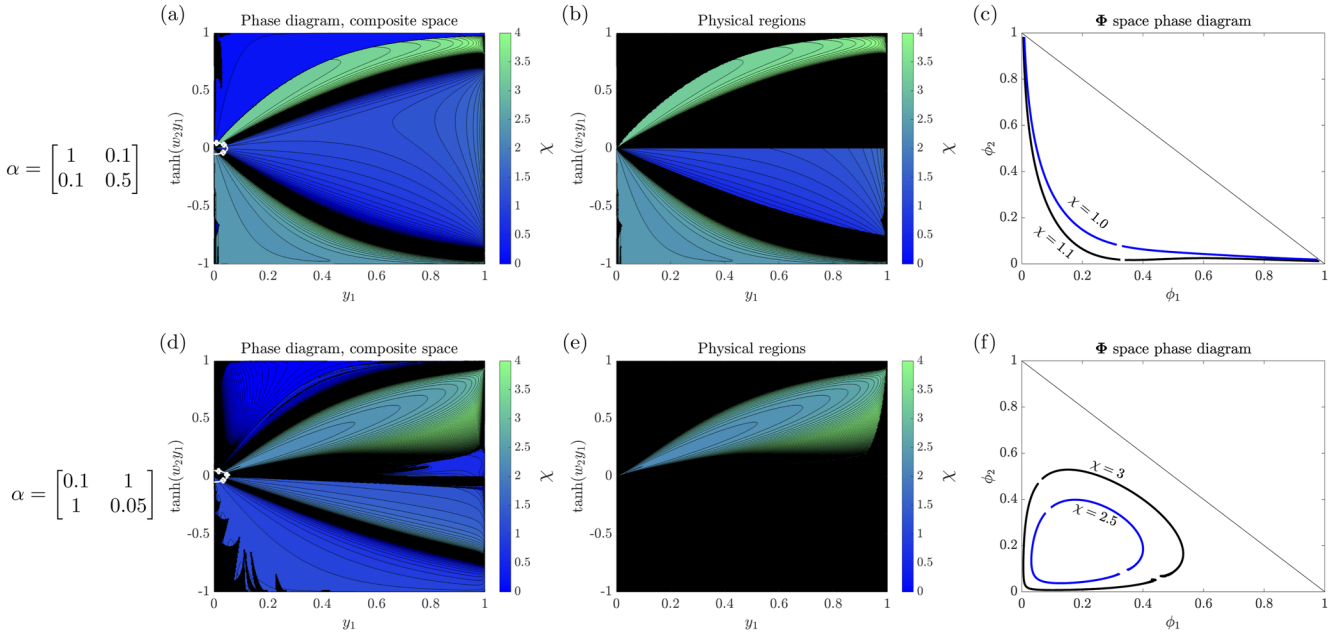

FIG. 4. Two-polymer–one-solvent system. Panels (a)–(c) correspond to the phase diagram for a diagonally dominant interaction matrix, while panels (d)–(f) show the same results for an off-diagonally dominant interaction matrix. (a) The full phase diagram mapping the χ surface (with colormap corresponding to the value of χ) for all compositional space for fixed α_{ij} on the y_1, $\tanh(y_1 w_2)$ plane. All the regions connect at the critical point at $y_1 = 0$ and $y_1 w_2 = 0$, but are segmented by the singular curves on the χ surface that are evident by the thick black regions on the χ surface. The three poles are computed on a small circle near the critical point and marked with white markers. These poles can be used to demarcate the distinct differentiable regions of the χ surface. Note that the colormap is saturated at the bounds of the color bar, but the contour lines are linear between $\chi = -10$ and $\chi = 10$ in increments of 0.1, although negative values of χ do not appear in the feasible physical regions. The black regions correspond to regions where no solution is available or where $|\chi| > 10$. (b) The regions of the χ surface that are physical, with positive predicted polymer densities. (c) The binodal curve for fixed values of χ transformed to $\mathbf{\Phi}$ space, either mapped out by the gradient descent algorithm in Sec. III C or by extracting contours from the physical regions of the χ surface in part 4(b). The small gaps correspond to the critical points of the binodal curves, which map to the same point in the composite composition coordinates. (d)–(f) The same information as panels (a)–(c), but for a different interaction matrix.

here are of limited use for generating single binodal curves at fixed χ.

At this point, it is natural to ask the following question: how can one efficiently sample the implicit χ surface to find a desired contour or sets of contours corresponding to multiple coexisting phases? Furthermore, how can one choose sampling points so as to dodge the singular curves in the phase diagram? Finally, can the sampling be efficiently scaled for large values of M? To explore these questions, in Sec. III D 3, we formulate such a sampling strategy.

3. Finding contours in a high dimensional phase diagram

In the case of a high dimensional multicomponent solution, we have found an implicit solution for χ, requiring only one nonlinear solve in one variable, z, to construct the surface,

$$\chi = \chi(y_1, \tanh(y_1 w_i)) = \chi(\mathbf{v}), \tag{66}$$

where only the M independent variables are listed as arguments in the vector $\mathbf{v}$.

Without enumerating all the possible inputs for the implicit function χ, we may not know beforehand where a given contour lies on the χ surface corresponding to coexisting phases. If we wish to only find the contours of this surface at fixed $\chi = \chi_0$ with minimal number of χ evaluations in a high dimensional phase diagram ($M \gg 1$), then we can employ a gradient descent algorithm on the χ surface to arrive at the desired contour. Defining the inputs of the χ function as a vector $\mathbf{v}$, with step index k, the algorithm proceeds as

$$\mathbf{v}_{k+1} = \mathbf{v}_k - \lambda_k \nabla_v F(\mathbf{v}). \tag{67}$$

The function F is

$$F(\mathbf{v}) = (\chi(\mathbf{v}) - \chi_0)^2, \tag{68}$$

and its gradient can be computed by a finite difference. The step size λ_k can be specified by a given algorithm, such as the Barzilai–Borwein method.[39]

Once the desired contour is reached, one can then trace out a contour by stepping along a direction, $\mathbf{u}$, perpendicular to the gradient $\mathbf{g} = \nabla_v \chi$. However, it should be noted that there are $M - 1$, such coordinate directions that are perpendicular to $\mathbf{g}$. Here, we choose just one of these coordinate directions,

$$\mathbf{v}_{k+1} = \mathbf{v}_k - \lambda_k \mathbf{u} \tag{69}$$

with

$$\begin{aligned} u_i &= g_i, \qquad i > 1, \\ u_i &= -\frac{1}{g_1} \sum_{i>1} g_i^2, \qquad i = 1. \end{aligned} \tag{70}$$

Here, the step size λ_k can be scaled by the magnitude of the gradient vector or the location of the current iteration, $\mathbf{v}_k$. The direction of the step away or toward the global critical point can be controlled by the sign of λ_k. If ever the step size is too large and the algorithm steps off of the fixed χ contour, the gradient descent can be used to return to the contour.

In a high dimensional composition space, the binodal curves become higher dimensional surfaces. Stepping in one tangential coordinate direction may not sample the entire binodal surface on its own. Therefore, one would need to choose a path that efficiently maps the boundary of the binodal surface and chooses the tangential directions based on the distance from the critical point, the attributes of the gradient vector, and the historical location of points already sampled. Here, we do not derive such a high dimensional search algorithm.

The gradient descent algorithm works to converge to one contour in the $\mathbf{v}$ space, corresponding to phase coexistence points in the $\mathbf{\Phi}$ space. If the algorithm is initiated in a region that is separated from the desired contour by a singular, non-differentiable curve in the $\mathbf{v}$ space, then more than one starting point would be needed to locate all matching contours in the phase diagram. If this is the case, then it is natural to ask: what are efficient ways to choose candidate starting points?

Notably, all the regions connect at the global critical point at $\mathbf{v} = 0$, but the regions are separated by singularities that extend radially from the critical point, slicing the phase diagram like a pie, as shown clearly in Figs. 4(a) and 4(d). Therefore, the analytical approximations of the equations by expanding near the critical point and finding poles can give a small number of candidate starting points in each region located at the center of each pie piece.

To sketch out the strategy to initiate a gradient descent algorithm, we next turn to how to segment the phase diagram efficiently to generate candidate starting points near the critical point.

a. Segmenting a hypersphere near the critical point. Near the critical point, the partitioning of each species becomes negligible, i.e., $|y_i|, |z| \ll 1$, and the master equation is

$$z^2 = \frac{\sum_i w_i y_i^2 / N_i}{\sum_i w_i} \tag{71}$$

with the chemical potential constraint setting,

$$y_i = N_i(\eta_i - 1)z + \eta_i \frac{N_i}{N_1} y_1. \tag{72}$$

Combined, these equations give a quadratic equation for z, which can be solved exactly, as documented in Appendix B. Nevertheless, extra degrees of freedom are still specified by the set y_1 and w_i.

To eliminate these degrees of freedom, we propose the following strategy: we look for candidate starting points on a small hypersphere $|\mathbf{v}| = v_0$ near the critical point, with $v_0 \ll 1/\max(N_i)$. First, we find the poles of the χ function on this hypersphere. Then, we choose a small number of points in regions demarcated by the poles of the χ function. By choosing specified points between the poles, we guarantee that we are testing each candidate region for the specified χ value.

For finite but small z and y_i, the poles of the χ function are defined by the equation,

$$\gamma_i \propto \sum_j \alpha_{ij} w_j = 0, \tag{73}$$

which corresponds to finding the null space of the $\boldsymbol{\alpha}$ matrix, and then rescaling so that $w_1 = 1$. If $\boldsymbol{\alpha}$ is full rank, then the null space is empty. However, one other denominator in the full expressions can go to zero, causing singularities in the phase diagram, namely,

$$\sum_j \left[w_j \left(1 + \frac{z}{y_j} \right) \right] = 0. \tag{74}$$

Near the critical point, these poles must be numerically solved for the values of w_i and y_1 that lead to singularities, making use of the exact solution for z near the critical point.

Note that the poles can actually be interpreted as curves on a given hypersphere. For $M = 2$, they correspond to points on a circle. For $M = 3$, they correspond to arcs on a sphere, and so on. In order to segment the space in higher dimensions, one would need to find the intersection of these singular curves on the hypersphere near the critical point, to find the boundaries of the nonsingular "continents" on the hypersphere.

Once we have found the poles near the critical point at fixed v_0, we can segment the hypersphere radially between these poles, making sure to test at least one point in the gradient descent algorithm in each "continent" of the segmented hypersphere. This would ensure that our algorithm can find a desired contour, if it exists, and could find multiple contours of the same function that are separated by non-differentiable curves.

b. Applying the gradient descent algorithm. Here, we explicitly apply the gradient descent algorithm only for $M = 2$, to find the contours shown in Figs. 4(c) and 4(f), validating the results by comparison to the full χ surface. The poles of the function do segment the phase diagram for $M = 2$, as demonstrated by the white points marked in Figs. 4(a) and 4(d) on a circle at a fixed distance from the critical point. Choosing a handful of starting values between the singular points on the circle, we can trace out the fixed values of χ, and then transform to $\mathbf{\Phi}$ space, to identify the curves matching shown in Figs. 4(c) and 4(f).

While we demonstrate the gradient descent algorithm for $M = 2$, we leave the exploration of $M > 2$ to future work. Note that this gradient descent algorithm is very advantageous for cases where $M > 4$, where it may be difficult to discretize the $\mathbf{v}$ space over the whole compositional range *a priori*.[27] In fact, while the number of function evaluations at each point scales with M (to evaluate the components of the gradient vectors), and the number of candidate points may also increase with M, the gradient descent algorithm is otherwise insensitive to M, which is much more favorable than discretizing all combinatoric possibilities in the composition space. Still, the stepping along the binodal surfaces in high dimensions must resolve enough points so as to map out the coexistance surface to a significant resolution, which may be a computationally demanding task.

IV. CONCLUSIONS

Using the implicit substitution method, we are able to analytically define the coexistence curves for a polymer–solvent system and a polydisperse polymer-solvent system. For generalized mixtures of many components, we have simplified a large system of nonlinear equations into one nonlinear equation in one composite variable. We have presented methods to approximate the solutions, not only for the single-polymer solvent system (Sec. II C) but also in the case of multicomponent solutions far from their critical point (Appendix A) and close to their critical point (Appendix B). More analytical solutions are possible in arbitrary mixtures if there is one dominant interaction, where all other non-interacting species operate as non-interacting crowders (Appendixes C and D).

As a test of the multicomponent master equation, we have mapped out the implicit χ surface for two polymeric components over the full compositional space and demonstrated the demarcation between differentiable regions of the χ surface that defines the binodal curve. Furthermore, we have suggested strategies to find full binodal curves at fixed χ without discretizing the full compositional space, which we believe will be very useful for $M > 4$, but we leave a full exploration of these methods to future studies.

Ample opportunities exist to further develop the methods presented here. Beyond improving the computational performance of the method for many components, one could seek analytical perturbative series solutions to the general master equation. Furthermore, one could implement the method to large multicomponent mixtures and explore the shape of contours of the χ surface for different interaction matrices. Importantly, the method could be applied to efficiently locate compositional regions of multiphase coexistence for large mixtures of components. Moreover, the approach can be directly extended to more detailed equations of state beyond Flory–Huggins, although compact analytical solutions of a final master equation are not guaranteed for all equations of state. In Appendix E, we highlight possible extensions of the implicit substitution method to other equations of state. Further, the method could be directly linked to the thermoresponsive behavior of polymer solutions using a model of the temperature dependence of χ.[40] The analysis applied here may also be useful for multicomponent regular solution models, now finding use even for phase transformations in battery materials.[41]

The composite composition variables are experimentally accessible if the composition of the dilute and concentrated phases is simultaneously measured. Therefore, experimental phase diagrams may also be mapped to composite composition space for direct comparison with the theoretical χ surface for a particular $\boldsymbol{\alpha}$ matrix at fixed temperature.

The implicit substitution method, while possibly un-intuitive at first sight, has given us a new lens to explore the phase behavior of polymeric phase separation. Here, we demonstrated its power in simplifying systems of coupled nonlinear thermodynamic equations that are often impossible to solve explicitly.

ACKNOWLEDGMENTS

We acknowledge support from the Princeton Center for Complex Materials, an NSF MRSEC (Grant No. DMR-2011750) and NSF Grant No. DMS/NIGMS 2245850. J.P.D. is also supported by the Omenn-Darling Princeton Bioengineering Institute–Innovators (PBI2) Postdoctoral Fellowship. We thank Hongbo Zhao, Tejas Dethe, and Qiwei Yu for helpful discussions and comments.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

J. Pedro de Souza: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal). **Howard A. Stone**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: APPROXIMATIONS FAR FROM THE CRITICAL POINT FOR MULTICOMPONENT MIXTURES

For multicomponent mixtures, it may be difficult to seek approximations, since multiple assumptions of limiting values of densities may need to be made simultaneously. While the limiting solution near the critical point is analytical (by applying the quadratic formula in Eq. (71) as detailed in Appendix B), one general approximation does not work in all the regions far from the critical point.

However, we can explore the limiting behavior when the major component, which we arbitrarily specify as component 1, has a large excess in the condensed phase compared to the dilute phase relative to other components. We also choose a secondary component that exhibits the opposite partitioning compared to component 1, but exhibits more partitioning compared to the other species. Therefore, this boundary corresponds to the edge of the phase diagram in our chosen coordinates, where $y_1 \to 1$, $w_i y_1 \to \pm\infty$, or $z \to 1$.

In our first analysis, we start with the solvent as the secondary dominant component. If this is the case, $y_1 \to 1-\epsilon$, where ϵ is vanishingly small, with a solvent dominated dilute phase, also meaning that $z \to 1$. We can effectively solve the system as if no other polymers are present. In the limit of $\epsilon \to 0$, where correspondingly w_i for $i > 1$ go to zero, we get

$$h(z) \approx 1 + (h(y_1) - 1)/N_1, \tag{A1}$$

which has solutions in terms of the inverse FH function,

$$z = h^{-1}(1 + (h(y_1) - 1)/N_1). \tag{A2}$$

We can subsequently solve for the set y_i describing the partitioning of other polymers by applying Eq. (56). We can map this approximation to all of composition space for a specific value of y_1 and w_i, and then check that the master equation is satisfied up to some tolerance to accept the approximation prediction or to reject it.

In many cases, although, the dilute phase in one polymer is the dense phase in the other, and vice versa, for a given polymer pair. In other words, the solvent may not be significantly partitioned between phases. If that is the case, we can seek approximations by assuming that one of the polymers is the "solvent" while the other is the polymer, effectively rescaling the theory in these dominant composition coordinates. If we were to do this with component j as the solvent and component y_i the dominant partitioning component, we would find

$$y_j = -h^{-1}(1 + (h(y_i) - 1)N_j/N_i), \tag{A3}$$

where y_j is chosen to have the opposite sign of y_i. All the other composite variables, y_k, can be computed by the analog of Eq. (56) that treats component j as the "solvent" component that satisfies incompressibility, namely,

$$y_k = \tanh\left(-\frac{N_k}{N_j}\left(\frac{\eta_k}{\eta_i} - 1\right)\tanh^{-1}(y_j) + \frac{\eta_k N_k}{\eta_i N_i}\tanh^{-1}(y_i)\right), \tag{A4}$$

where for the solvent $y_0 = -z$.

We can use this procedure to map out the areas near the boundary of our composition space, where $y_1 \to 1$ or $w_i y_1 \to \pm\infty$. Furthermore, by considering each possible pair of polymers or solvent combination, we can generate a combinatoric set of approximations that span all regions far from the critical point, which may at least make for useful guesses for the full master equation. Similar to the one-polymer case, we expect that these approximations will be most accurate when there are significant differences in the polymer lengths.

APPENDIX B: APPROXIMATIONS NEAR THE CRITICAL POINT

Near the critical point, the partitioning of each species becomes negligible, i.e., $|y_i|, |z| \ll 1$, and the master equation is

$$z^2 = \frac{\sum_i w_i y_i^2/N_i}{\sum_i w_i}, \tag{B1}$$

with the chemical potential constraint setting,

$$y_i = N_i(\eta_i - 1)z + \eta_i \frac{N_i}{N_1} y_1. \tag{B2}$$

This approximation is valid as long as

$$\begin{aligned} z \ll 1 \quad \& \quad z \ll \frac{1}{N_i(\eta_i - 1)}, \\ y_1 \ll 1 \quad \& \quad y_1 \ll \frac{N_1}{N_i \eta_i}. \end{aligned} \tag{B3}$$

Combined, these equations give a quadratic equation for z, which can be solved exactly. The quadratic equation explicitly is

$$\begin{aligned} 0 = z^2\left(-1 + \frac{\sum_i w_i N_i(\eta_i - 1)^2}{\sum_i w_i}\right) \\ + z\left(\frac{\sum_i 2 w_i N_i \eta_i (\eta_i - 1) y_1/N_1}{\sum_i w_i}\right) \\ + \frac{\sum_i w_i N_i \eta_i^2 y_1^2/N_1^2}{\sum_i w_i}. \end{aligned} \tag{B4}$$

Picking the set w_i and y_1, we solve for z, choosing the root that gives positive β_i for all i, if, in fact, either root gives positive β_i.

APPENDIX C: ADDING NON-INTERACTING "CROWDERS"

To add to our list of analytical solutions, we can consider the case of non-interacting crowders added to a single strongly interacting polymer in solvent. Although this ideal non-interacting assumption may be rare in reality, it may be a reasonable approximation if one polymer self-interaction is much larger in magnitude than all other interactions. We assume that the dominant polymer is $i = 1$, where $\alpha_{11} = 1$. If non-interacting crowders are added with species index: $2 < i < M$, then since their corresponding rows and columns $\alpha_{ij} = 0$, we have $\eta_i = 0$ for these species. Therefore, the $\boldsymbol{\alpha}$ matrix only has one entry.

For the non-interacting polymers, we have

$$y_i = -\tanh\left(N_i \tanh^{-1}(z)\right). \tag{C1}$$

The master equation can be written as

$$h(z) - 1 = \frac{(h(y_1) - 1)/N_1}{\sum_i w_i} + \frac{\sum_{i=2}^{M} w_i \left(\frac{N_i \tanh^{-1}(z)}{\tanh\left(N_i \tanh^{-1}(z)\right)} - 1\right)/N_i}{\sum_i w_i}. \tag{C2}$$

While z is embedded in multiple nonlinear functions, y_1 appears as only one argument of a nonlinear function. Therefore, instead of specifying y_1 and solving for z, we can specify z and solve for y_1,

$$h(y_1) = 1 + N_1 \sum_i w_i (h(z) - 1) + \sum_{i=2}^{M} w_i \left(1 - \frac{N_i \tanh^{-1}(z)}{\tanh\left(N_i \tanh^{-1}(z)\right)}\right)\frac{N_1}{N_i}. \tag{C3}$$

In this case, y_1 can explicitly be expressed in terms of the inverse FH function.

Similar to the polydisperse problem, we may know the overall ratios of the crowders being added to the solution relative to the total solvent being added. We can express this ratio in terms of z and v (the volume fraction of the dense phase A),

$$\begin{aligned}\tilde{\phi}_i &= v\phi_{iA} + (1 - v)\phi_{iB} = (\phi_{iA} - \phi_{iB})\left(v + \frac{1}{2y_i} - \frac{1}{2}\right) \\ &= (\phi_{iA} - \phi_{iB})\left(v - \frac{1}{2\tanh\left(N_i \tanh^{-1}(z)\right)} - \frac{1}{2}\right),\end{aligned} \tag{C4}$$

then the ratio between the average polymer density of component i and the solvent $\tilde{\phi}_i/(1 - \sum_j \tilde{\phi}_j)$ is

$$\frac{\tilde{\phi}_i}{1 - \sum_j \tilde{\phi}_j} = \frac{\left(v - \frac{1}{2\tanh\left(N_i \tanh^{-1}(z)\right)} - \frac{1}{2}\right)}{\left(-v + \frac{1}{2z} + \frac{1}{2}\right)} \frac{w_i}{\sum_j w_j}. \tag{C5}$$

If the left-hand side of the equation (the ratio the crowder to solvent) is known for $i > 1$, and v and z are specified, then the above-mentioned equations give $M - 1$ linear equations for the $M - 1$ values of w_i for $i > 1$.

Therefore, if v and z are specified with a known total composition of the crowders and solvent, then the set y_1 and w_i are all specified. To construct a phase diagram with crowders present, we can simply vary v and z between 0 and 1 to construct the full χ surface and then transform to $\boldsymbol{\Phi}$ space.

It should be noted that we may also use this "non-interacting" construction to include finite "compressibility" of our solution, by including "holes" on the lattice as one of the non-interacting species, along with the solvent and polymer species. However, this lattice-based approximation is typically a poor approximation for the compressibility of hard-sphere liquids[42] and, therefore, can likely only approximate the behavior of real liquids.

APPENDIX D: ONE DOMINANT POLYMER–POLYMER INTERACTION

Analogous to the case of one dominant self-interacting polymer is the case of one dominant polymer–polymer pair interaction, meaning an interaction matrix that is zero everywhere except for two off-diagonal entries. If we are to reframe the theory such that one of the polymers is made into the "solvent" species and the other is made into the "dominant" interacting component, then the approximation becomes equivalent to Appendix C, where all other species are only non-interacting crowders that are non-participating observers to the main interacting components. In these reframed variables, the interaction matrix would have only one non-zero entry, α_{11}.

Note that for the case where the "solvent" species has length $N_0 > 1$, we can rewrite the master equation as

$$h(z) = 1 + \frac{\sum_i w_i (h(y_i) - 1) N_0/N_i}{\sum_i w_i} \tag{D1}$$

and the chemical potential constraints give y_i

$$y_i = \tanh\left(N_i(\eta_i - 1)\tanh^{-1}(z)/N_0 + \eta_i N_i \tanh^{-1}(y_1)/N_1\right). \tag{D2}$$

Next, we could pursue the same approximations as in the previous sections, where $\eta_i = 0$ for $i > 1$ and then frame the phase diagram in terms of v and z. We leave an exploration of these approximations to future work.

It should be noted that there is a subtle difference between the "dominant-interaction-approximation" in this section as compared to the "far from critical" approximation in Appendix A. The dominant-interaction approximation is valid only when the set $\eta_i \ll 1$ for $i \neq 1$, regardless of the geometric distance from the critical point. However, for sufficiently strong partitioning of a weakly interacting component k, $|w_k| \gg 1$, the condition on η_k will eventually break. So, in other words, the "dominant-interaction approximation" breaks very far from criticality for weakly interacting components. The "far from critical" approximation, on the other hand, is even valid when the interaction matrix has many non-zero entries, although its validity is only maintained far from the critical point.

To summarize all the approximations explored, building from Flory's original approach for a polymer–solvent system, we have derived an analytical approximation far from the critical point. Next, we have derived an approximation valid in a small region near the critical point. To bridge these approximations, we have explored regions of the phase diagram where a particular polymer–polymer or polymer–solvent interaction is dominant. While more intricate approximations may be sought for triplets of species, we suspect that for most regions in the phase diagram, the phase composition is dominated by the two main interacting accumulated/depleted components between the phases, and if not, we are either close to the

global critical point or very far from the global critical point—where other approximations are more readily accessible.

APPENDIX E: EXTENSIONS TO OTHER EQUATIONS OF STATE

Here, we can explore how this formulation could be extended to other equations of state based on a lattice, focusing only on the polymer–solvent system. In other theories, one could replace the $-\chi\phi^2$ term in the free energy density by

$$-\chi\tilde{f}_{\text{int}}(\phi), \tag{E1}$$

which is equivalent to assuming a ϕ-dependent χ parameter.

In this case, from the chemical potential, we arrive at

$$\chi = \frac{\frac{2}{N}\tanh^{-1}(y) + 2\tanh^{-1}(z)}{\tilde{f}'_{\text{int}}(\phi_A) - \tilde{f}'_{\text{int}}(\phi_B)}, \tag{E2}$$

and from the pressure, we have

$$\chi = \frac{\left(\frac{1}{N} - 1\right)(\phi_A + \phi_B)y + 2\tanh^{-1}(z)}{\tilde{f}_{\text{int}}(\phi_B) - \tilde{f}_{\text{int}}(\phi_A) + \phi_A\tilde{f}'_{\text{int}}(\phi_A) - \phi_B\tilde{f}'_{\text{int}}(\phi_B)}. \tag{E3}$$

Substituting in

$$\begin{aligned}\phi_A &= \frac{z(1+y)}{z+y},\\ \phi_B &= \frac{z(1-y)}{z+y},\end{aligned} \tag{E4}$$

and equating the values of χ, we generally have a nonlinear equation for one unknown (z) in terms of one specified variable (y)—reducing our system of two equations down to one equation. While not guaranteed, one may be able to decouple the z and y dependence, defining a function similar to the FH function $h()$ for a given functional interaction form. While this procedure is marginally better than solving the full system of two equations for the polymer–solvent system, it may be especially helpful for the multicomponent case. We leave these generalizations to future work.

REFERENCES

[1] M. L. Huggins, J. Chem. Phys. **9**, 440 (1941).
[2] P. J. Flory, J. Chem. Phys. **10**, 51 (1942).
[3] M. J. Cantow, *Polymer Fractionation* (Academic Press, Inc., 1967).
[4] A. F. Barton, *Handbook of Poylmer-Liquid Interaction Parameters and Solubility Parameters* (Routledge, 2018).
[5] K. Milczewska and A. Voelkel, J. Polym. Sci., Part B: Polym. Phys. **44**, 1853 (2006).
[6] K.-V. Peinemann, V. Abetz, and P. F. Simon, Nat. Mater. **6**, 992 (2007).
[7] P. P. Angelopoulou, L. T. Kearney, J. K. Keum, L. Collins, R. Kumar, G. Sakellariou, R. C. Advincula, J. W. Mays, and K. Hong, J. Mater. Chem. A **11**, 9846 (2023).
[8] C. P. Callaway, K. Hendrickson, N. Bond, S. M. Lee, P. Sood, and S. S. Jang, ChemPhysChem **19**, 1655 (2018).
[9] C. P. Brangwynne, P. Tompa, and R. V. Pappu, Nat. Phys. **11**, 899 (2015).
[10] J. Berry, C. P. Brangwynne, and M. Haataja, Rep. Prog. Phys. **81**, 046601 (2018).
[11] R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, and M. Kar, Chem. Rev. **123**, 8945 (2023).
[12] J. T. Overbeek and M. J. Voorn, J. Cell. Comp. Physiol. **49**, 7 (1957).
[13] R. Simha and T. Somcynsky, Macromolecules **2**, 342 (1969).
[14] J. Scheutjens and G. Fleer, J. Phys. Chem. **84**, 178 (1980).
[15] F. Tanaka, Macromolecules **22**, 1988 (1989).
[16] F. Tanaka and W. H. Stockmayer, Macromolecules **27**, 3943 (1994).
[17] A. N. Semenov and M. Rubinstein, Macromolecules **31**, 1373 (1998).
[18] A.-V. G. Ruzette and A. M. Mayes, Macromolecules **34**, 1894 (2001).
[19] Z.-G. Wang, J. Phys. Chem. B **112**, 16205 (2008).
[20] Z.-G. Wang, Macromolecules **50**, 9073 (2017).
[21] J. Wessén, T. Pal, S. Das, Y.-H. Lin, and H. S. Chan, J. Phys. Chem. B **125**, 4337 (2021).
[22] I. C. Sanchez, Macromolecules **17**, 967 (1984).
[23] C. J. Tsenoglou and C. D. Papaspyrides, Polymer **42**, 8069 (2001).
[24] S. Scheinhardt-Engels, F. Leermakers, and G. Fleer, Phys. Rev. E **69**, 021808 (2004).
[25] D. Qian, T. C. Michaels, and T. P. Knowles, J. Phys. Chem. Lett. **13**, 7853 (2022).
[26] S. H. van Leuken, R. A. van Benthem, R. Tuinier, and M. Vis, Macromol. Theory Simul. **32**, 2300001 (2023).
[27] S. Mao, D. Kuldinow, M. P. Haataja, and A. Košmrlj, Soft Matter **15**, 1297 (2019).
[28] W. M. Jacobs and D. Frenkel, Biophys. J. **112**, 683 (2017).
[29] W. M. Jacobs, J. Chem. Theory Comput. **19**, 3429 (2023).
[30] M. L. Huggins and H. Okamoto, *Polymer Fractionation* (Elsevier, 1967), pp. 1–42.
[31] M. Wühr, R. M. Freeman, M. Presler, M. E. Horb, L. Peshkin, S. P. Gygi, and M. W. Kirschner, Curr. Biol. **24**, 1467 (2014).
[32] M. Doi, *Soft Matter Physics* (Oxford University Press, 2013).
[33] M. Bazant, Electrochemical Energy Systems, Spring 2014: Lecture 11, 2014, https://ocw.mit.edu/courses/10-626-electrochemical-energy-systems-spring-2014/resources/mit10_626s14_s11lec11/.
[34] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).
[35] J. P. de Souza and H. A. Stone (2024). "FH-Binodal," Github. https://github.com/jpldesouza/FH-Binodal
[36] S. H. van Leuken, R. A. van Benthem, R. Tuinier, and M. Vis, J. Phys.: Mater. **7**, 015005 (2023).
[37] D. Blankschtein, *Lectures in Classical Thermodynamics with an Introduction to Statistical Mechanics* (Springer, 2020), Vol. 766.
[38] A. Bot, E. van der Linden, and P. Venema, ACS Omega **9**, 22677 (2024).
[39] J. Barzilai and J. M. Borwein, IMA J. Numer. Anal. **8**, 141 (1988).
[40] S. Dhamankar and M. A. Webb, "Asymmetry in polymer–solvent interactions yields complex thermoresponsive behavior," ACS Macro Lett. **13**, 818–825 (2024).
[41] R. B. Smith, E. Khoo, and M. Z. Bazant, "Intercalation kinetics in multiphase-layered materials," J. Phys. Chem. C **121**(23), 12505–12523 (2017).
[42] D. Frydel and Y. Levin, J. Chem. Phys. **137**, 164703 (2012).