

# Supplementary Information: Physical theory of biological noise buffering by multicomponent phase separation

Dan Deviri <sup>\*1</sup> and Samuel A. Safran<sup>1</sup>

<sup>1</sup>Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 76100, Israel

## 1 Continuum theory of non-equilibrium LLPS

The main and novel claim of our paper is that noise buffering is most effective when the noise is correlated such that the distribution of overall solute concentrations is aligned with the directions of the tie-lines. We arrived at this conclusion by considering the superposition of the “noise ellipses” and the equilibrium phase diagrams, which describe the coexisting concentrations of solutes within the two phases. In this section, we justify this superposition from the separation of time scales of relatively slowly fluctuating protein expression, slow protein degradation, and the relatively fast molecular diffusion along gradients of the chemical potential; the latter is related to the tendency for free energy minimization. We demonstrate this from a general theory for the non-equilibrium dynamics of the system, that considers on the same footing, both the expression noise and chemical potential-driven molecular diffusion (which leads to liquid-liquid phase separation, LLPS). In the absence of expression and degradation of proteins, the system would be characterized by equilibrium LLPS. In the presence of slowly varying expression and degradation, the equilibrium phase diagram still characterizes the coexisting concentrations, but the non-equilibrium production and degradation of the proteins can affect the sizes and shapes of the phase separated domains – which for an equilibrium system would be of order of the system size. This validates our predictions based on superposition of the noise ellipses and the equilibrium phase diagrams.

We begin by formulating the dynamical equations that determine the concentrations of the two solutes as functions of time and position,  $\phi(\vec{r}, t)$  and  $\psi(\vec{r}, t)$ . In the cell, proteins and RNA molecules are expressed from genes and eventually degrade. During this time, these biological molecules diffuse along their chemical potential gradients. Therefore, the equations for the time evolution of the concentrations  $\phi(\vec{r}, t)$  and  $\psi(\vec{r}, t)$ , combine continuity equations with production and degradation kinetics.

$$\frac{d\phi(\vec{r}, t)}{dt} = \Gamma_\phi(\vec{r}, t) - k_d^\phi \phi(\vec{r}, t) - \nabla \cdot \vec{J}_\phi \quad (1)$$

$$\frac{d\psi(\vec{r}, t)}{dt} = \Gamma_\psi(\vec{r}, t) - k_d^\psi \psi(\vec{r}, t) - \nabla \cdot \vec{J}_\psi \quad (2)$$

where  $\Gamma_\phi(\vec{r}, t)$  and  $\Gamma_\psi(\vec{r}, t)$  are respectively the time and space-dependent production rates of solutes  $\phi$  and  $\psi$ ,  $k_d^\phi$  and  $k_d^\psi$  are respectively the degradation rates of solutes  $\phi$  and  $\psi$ , and  $\vec{J}_\phi$  and  $\vec{J}_\psi$  are respectively the fluxes of solutes  $\phi$  and  $\psi$ . We note that in principle, the degradation rates can also vary in time and with position. However, for simplicity, we approximate them as constants. This simplification is justified biologically for situations in which, for example, the enzymes that mark proteins for degradation by ubiquitination are homogeneously distributed within the cytoplasm.

---

\*Corresponding author: dan.deviri@weizmann.ac.il

We consider the common biological case where the solutes  $\phi$  and  $\psi$  are proteins or RNA molecules that cannot be exchanged with those comprising the extracellular environment. This implies that components of the fluxes  $\vec{J}_\phi$  and  $\vec{J}_\psi$  perpendicular to the boundaries of the system must vanish. Therefore, when Eqs. 1 and 2 are integrated over the volume of the system, the integrals over the divergences of the fluxes vanish, as imposed by the divergence theorem. This leads to the following pair of equations that describe the overall, system-averaged concentrations  $\bar{\phi}(t) = (\int \phi(\vec{r}, t) d\vec{r})/V$  and  $\bar{\psi}(t) = (\int \psi(\vec{r}, t) d\vec{r})/V$  (where  $V$  is the volume of the system):

$$\frac{d\bar{\phi}(t)}{dt} = \bar{\Gamma}_\phi(t) - k_d^\phi \bar{\phi}(t) \quad (3)$$

$$\frac{d\bar{\psi}(t)}{dt} = \bar{\Gamma}_\psi(t) - k_d^\psi \bar{\psi}(t) \quad (4)$$

where  $\bar{\Gamma}_\phi(t) \equiv (\int \Gamma_\phi(\vec{r}, t) d\vec{r})/V$  and  $\bar{\Gamma}_\psi(t) \equiv (\int \Gamma_\psi(\vec{r}, t) d\vec{r})/V$  are respectively the spatially averaged production rates of solutes  $\phi$  and  $\psi$ . It is important to note that the average production rates  $\bar{\Gamma}_\phi$  and  $\bar{\Gamma}_\psi$  are, in general, stochastic. This gives rise to the “noise ellipses”, which we derive and discuss in subsection “Theoretical derivation of the shape and orientation of the noise ellipse” below.

As can be seen, Eqs. 3 and 4, which determine the dynamics of the overall concentrations of the solutes, involve only production and degradation, and not the molecular interactions and entropy of the solutes which give rise to LLPS; these are included in the fluxes  $\vec{J}_\phi$  and  $\vec{J}_\psi$  in Eqs. 1 and 2. The derivation of Eqs. 3 and 4, which determine the noise ellipses, from Eqs. 1 and 2, which determine the  $\phi - \psi$  phase diagram and tie-lines, provides some justification of the superposition of the noise-ellipse on the phase diagrams. In the following subsections, we use these two pairs of equations as starting points to theoretically validate our theory in more detail and derive two important results used in our work. The first, which we derive from Eqs. 1 and 2, is that in the limit of fast diffusion relative to degradation, which is appropriate for biological systems, equilibrium phase diagrams indeed predict the coexisting concentrations of cellular LLPS. This is true despite the out-of-equilibrium nature of the system due to protein production and degradation. The second result, which we derive from Eqs. 3 and 4 using linear response theory, is the ellipsoidal shape of the distribution of the (noisy) overall solute concentrations when plotted in the  $\phi - \psi$  diagram.

## 1.1 Theory of non-equilibrium LLPS

The fluxes of the solutes that appear in Eqs. 1 and 2,  $\vec{J}_\phi$  and  $\vec{J}_\psi$ , are driven by gradients of their respective chemical potentials,  $\mu_\phi$  and  $\mu_\psi$ . These chemical potentials are related to the total free energy functional  $F$  of the system (that accounts for the entropy and interaction energies) by the functional derivatives,  $\mu_\phi = \delta F/\delta\phi$  and  $\mu_\psi = \delta F/\delta\psi$ . To linear order, the fluxes are related to the gradients of the chemical potentials by  $\vec{J}_i = -D_i \nabla \mu_i/k_B T$ , where  $i = \phi, \psi$ ,  $k_B T$  is the thermal energy and  $D_i$  is the mobility of solute  $i$ , which has the dimensions of Length<sup>5</sup>/Time and is related to the diffusion coefficient [1]. In general, the mobilities can be depend on the volume fractions  $\phi, \psi$  and their gradients (and in cellular systems, all other solutes that are present). However, we are interested in the limit where the primary driving force for the motion of particles is associated with gradients of the chemical potentials. Consistent with many other treatments [1, 2], we approximate the mobilities by constants, so that variations of the fluxes of solutes are only caused by variations of their chemical potentials gradients.

With the approximation of constant mobilities, Eqs. 1 and 2 can be written as:

$$\frac{d\phi(\vec{r}, t)}{dt} = k_d^\phi (\Phi(\vec{r}, t) - \phi(\vec{r}, t)) + D_\phi \nabla^2 \left( \frac{\delta}{\delta\phi} \left( \frac{F}{k_B T} \right) \right) \quad (5)$$

$$\frac{d\psi(\vec{r}, t)}{dt} = k_d^\psi (\Psi(\vec{r}, t) - \psi(\vec{r}, t)) + D_\psi \nabla^2 \left( \frac{\delta}{\delta\psi} \left( \frac{F}{k_B T} \right) \right) \quad (6)$$

where  $\Phi = \Gamma_\phi/k_d^\phi$  and  $\Psi = \Gamma_\psi/k_d^\psi$ , are effective concentrations associated with the production and degradation kinetics. Their volumetric integrals, according to Eqs. 3 and 4, determine the steady-state values of the overall solute concentrations  $\bar{\phi}$  and  $\bar{\psi}$ .

The time for proteins to diffuse across eukaryotic cells, (related to the mobilities  $D_i$ ), is of the order of 10 seconds [3]. Diffusion times on the scale of LLPS domains are even shorter. This contrasts with the timescale associated with changes in production of proteins or RNA molecules or their degradation which are hours, as we now explain.

For proteins, the volume fractions  $\phi$  and  $\psi$  are determined by the expression of genes whose dynamics are commonly controlled by transcriptional bursts [4]. Transcriptional bursts are short, stochastic events in which gene transcription becomes active, leading to synthesis of multiple copies of mRNA molecules which are then translated by the ribosomes until they are degraded. This production mechanism implies that two timescales underlie the variation of the production terms  $\Phi(\vec{r}, t)$  and  $\Psi(\vec{r}, t)$ , the average time between bursts and the degradation time of the mRNA molecules [4]. Both of these timescales are of the order of hours in human cells [3, 4].

Therefore, a significant separation of timescales of three orders of magnitude exists between the timescale over which  $\Phi(\vec{r}, t)$  and  $\Psi(\vec{r}, t)$  vary, compared with the timescale over which the concentration profiles of the solutes,  $\phi(\vec{r}, t)$  and  $\psi(\vec{r}, t)$ , reach steady state. It is this separation of timescales which gives rise to the quasi steady-state nature of the system. The concentrations profile  $\phi(\vec{r}, t)$  and  $\psi(\vec{r}, t)$  change very slowly over time and can to a very good approximation, be determined by Eqs. 5 and 6 with their left hand sides set to zero [5]. Corrections to this approximation due to the temporal variations of the production terms are of the order of the ratio of the diffusion time and the production variation timescale, which is about  $10^{-3}$ . We note that this separation of timescales is still valid if solute  $\phi$ , solute  $\psi$ , or both are RNA molecules because the time between transcriptional bursts is much longer than the diffusion of the RNA molecules [3, 4].

Motivated by the analogy discussed by [6, 7], we rewrite Eqs. 5 and 6 in the quasi-static approximation (in which their left hand sides are approximately set to zero), as follows:

$$0 = \nabla^2 \left( \frac{\delta}{\delta\phi} \left( \frac{F}{k_B T} + \frac{k_d^\phi}{8\pi D_\phi} \int \frac{(\Phi(\vec{r}, t) - \phi(\vec{r}, t))(\Phi(\vec{r}', t) - \phi(\vec{r}', t))}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \right) \right) \quad (7)$$

$$0 = \nabla^2 \left( \frac{\delta}{\delta\psi} \left( \frac{F}{k_B T} + \frac{k_d^\psi}{8\pi D_\psi} \int \frac{(\Psi(\vec{r}, t) - \psi(\vec{r}, t))(\Psi(\vec{r}', t) - \psi(\vec{r}', t))}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \right) \right) \quad (8)$$

where we used the identity  $\nabla^2 (|\vec{r} - \vec{r}'|^{-1}) = -4\pi\delta(\vec{r} - \vec{r}')$ , with  $\delta(\vec{r})$  being the Dirac delta function. Eqs. 7 and 8 allow us to define an effective free energy functional  $\tilde{F}$  that includes contributions of the non-equilibrium terms:

$$\tilde{F} = F + \frac{k_B T k_d^\phi}{8\pi D_\phi} \int \frac{(\Phi(\vec{r}, t) - \phi(\vec{r}, t))(\Phi(\vec{r}', t) - \phi(\vec{r}', t))}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \quad (9)$$

$$+ \frac{k_B T k_d^\psi}{8\pi D_\psi} \int \frac{(\Psi(\vec{r}, t) - \psi(\vec{r}, t))(\Psi(\vec{r}', t) - \psi(\vec{r}', t))}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \quad (10)$$

Expressing the kinetic equations in terms of the effective free energy  $\tilde{F}$  allows us to rewrite Eqs. 7 and 8 simply as  $\nabla^2 (\delta\tilde{F}/\delta\phi) = 0$  and  $\nabla^2 (\delta\tilde{F}/\delta\psi) = 0$ , respectively. In isotropic systems, namely systems with no preferred direction, the only non-diverging solution to the equations  $\nabla^2 (\delta\tilde{F}/\delta\phi) = 0$  and  $\nabla^2 (\delta\tilde{F}/\delta\psi) = 0$ , is  $\delta\tilde{F}/\delta\phi = \tilde{\mu}_\phi$  and  $\delta\tilde{F}/\delta\psi = \tilde{\mu}_\psi$ , where  $\tilde{\mu}_\phi$  and  $\tilde{\mu}_\psi$  are constants. These equations are the ones that determine the volume fractions  $\phi(\vec{r})$  and  $\psi(\vec{r})$  in the equilibrium situation characterized by the free energy  $\tilde{F}$ . In summary, we have shown that non-equilibrium LLPS that is relevant to cellular environment can be mapped to an equilibrium LLPS with an effective free energy functional. This effective free energy functional is the standard free energy that describes the short-range molecular interactions and entropy of the molecules, supplemented by terms that represent effective long-range interactions between the phase separating solutes (due to production and degradation), which are analogous to electrostatic repulsion. The form of the effective long-rang interactions is similar to those that appear in theories of solutions with reacting components, known as active emulsions [8]. We now determine to what extent the the long-range interaction terms in the effective

free energy of Eq. 9, modify the nature of the non-equilibrium LLPS compared with the equilibrium case without production and degradation, given by minimization of the free energy  $F$ .

First, it is important to note that although the terms due to the non-equilibrium production and degradation in Eq. 9 represent long-range interactions, they are relatively weak. This is because the separation of time scales described above: the timescale of degradation of proteins and RNA molecules is of the order of hours while the diffusion of these molecules across the cell is of the order of ten seconds [3]. Therefore, the prefactors  $k_d^i / (8\pi D_i)$  are small so that the contributions of the non-equilibrium terms in Eq. 9 may be significant only over long length scales since they are then integrated and accumulated over large volumes. For this reason, the effect of these terms on the equilibrium LLPS properties induced by the entropy and short-range molecular interactions (that are included in the free energy  $F$  in Eq. 9) are not expected to change the local, small-scale properties of the LLPS, namely the coexisting concentrations within the different phases. Instead, these weak, long-range interactions can change large-scale properties such as the spatial organization of phase separated domains, including their number and shapes.

Second, the long-range terms in Eq. 9 are analogous [6, 7] to electrostatic Coulomb energy with effective “charge densities” of  $\rho_\phi(\vec{r}, t) \sim \Phi(\vec{r}, t) - \phi(\vec{r}, t)$  and  $\rho_\psi(\vec{r}, t) \sim \Psi(\vec{r}, t) - \psi(\vec{r}, t)$ . From this analogy, as we argue below, we expect that in steady state, the concentration profiles are constrained by the charge “neutrality” (here, corresponding to a certain average concentration), so that the signs of the effective “charge densities” alternate throughout the volume of the cell. The effective “attraction” of these alternating “charges” reduces the long-range terms in the effective free energy. Of course, the steady state is determined by minimization of the full, effective free energy in Eq. 9, including  $F$ , which increases in the case of spatial gradients of the concentrations due to gradient terms. Indeed, these give rise to the interfacial tensions of finite-size domains. Complete determination of the concentration profiles requires detailed information on the spatial dependence of the production, as well as on the molecular interactions that give rise to the free energy  $F$ . However, we note that a fundamental property of the “charge densities”,  $\rho_\phi(\vec{r}, t) \sim \Phi(\vec{r}, t) - \phi(\vec{r}, t)$  and  $\rho_\psi(\vec{r}, t) \sim \Psi(\vec{r}, t) - \psi(\vec{r}, t)$ , is that integration of these “charge densities” over the volume of the cell in steady state must vanish as implied by Eqs. 3 and 4. Therefore, there must be both positive and negative “effective charges” within the volume of the system. This motivates us to consider two limiting cases:

(1) Production is localized to a finite region, such as RNA that is transcribed from a single locus, as previously considered by Riback et al. [9]. This means that within this production region, the effective molecular concentrations,  $\Phi$  and  $\Psi$ , must be large enough to balance their integrated solute concentrations  $\phi$  and  $\psi$  over the entire volume of the system. Therefore, within the production region,  $\Phi$  and  $\Psi$  must be significantly larger than the local concentrations of their respective solutes. In that scenario, the local effective charge in the production region is positive and large. Outside of the production region the charge is, by definition, negative because it is proportional to the negative value of the volume fraction of the solute. We therefore expect that in the vicinity of the local production region with a very high positive, effective “charge”, a solute-rich phase will nucleate so that the negative charge outside this the production region is maximized to lower the effective “Coulomb” interaction energy. If there are multiple, localized production regions (e.g. RNA that is transcribed from multiple loci) the system is no longer spherically symmetric, giving rise to multiple, solute-rich phases or distorted ones with aspherical shapes. The shapes and number of solute-rich phases depend on the magnitude of the interfacial tension, ratio between diffusion and degradation times (which determines the strength of the effective long-range interactions), and the detailed positions and sizes the sources. This scenario is consistent with the model and observations of Riback et al. [9], who demonstrated that solute fluxes from multiple production regions are responsible for aspherical shapes of nucleoli in pathological conditions, as well as for other aspherical biological condensates.

(2) The production is relatively homogeneous and spatially-independent, e.g. by uniformly dispersed ribosomes in the cytoplasm. In that case, the “charge” in the solute-rich phases is negative and in the solute-poor regions is positive. This gives rise to alternating “charge” patterns of solute-rich and solute-poor domains. It is important to note that in the equilibrium problem without production and degradation, there are periodic solutions with any period that are local (but not global) minima of the free energy  $F$  [10, 11]. However, the free energy of such periodic solutions increases with the number of interfaces between dense and dilute domains. This interfacial tension effect means that in the equilibrium case, a single interface is the global minimum of the free energy. Interestingly, in the non-equilibrium case, the competition of

the interfacial tension free energy, which favors a single interface, with the long-range interactions due to the production and degradation terms, which favor multiple interfaces, results in an optimal domain size. Globally, this gives rise to a microstructure of alternating LLPS domains of finite size (ignoring fluctuations, a periodic structure), similar to patterns that form in active emulsions [8, 12]. This may explain the observed stability of states of multiple solute rich domains in the cytoplasm (e.g., in [13]). We now predict the scaling of the typical domain size which is the global minimum of  $\bar{F}$ , as a function of the ratios  $k_d^i/D_i$  in Eq. 9. For simplicity, we consider only one solute that is produced and degraded whose ratio of degradation rate and mobility is  $k_d/D$ ; this case can be easily generalized to one of two solutes.

A dimensional analysis of the effective free energy, which we present below, along with the physical requirement that it be extensive in the size of the system, shows that the first contribution  $F$  to the effective free energy in Eq. 9, scales as the volume of the system divided by the size of a typical phase separated domain. This is indeed the driving force for coarsening in the equilibrium system. On the other hand, a similar analysis shows that the other terms in Eq. 9, arising from protein production and degradation, and are analogous to “electrostatics”, scale as the system volume times the square of the domain size. From this analysis, the domain size that minimizes the effective free energy scales as the cube root of the ratio  $\sigma D/k_B T k_D$ . The same scaling can be predicted from a more quantitative calculation of a one-dimensional array of domains. These are indeed the minima of  $F$  for one-dimensional periodicity whose concentration profiles can be calculated and then used in perturbation theory, to calculate the “electrostatic” contributions. However, here, we use a more informal, but quantitative estimate for the effective free energy of a three-dimensional system with finite-sized domains.

We consider a periodic concentration profile in three dimensions described by a “lattice” made of domains of size  $L$ , where  $L$  is the period. For ease of calculation, we approximate these domains to be spheres of radius  $L/2$ , with a concentric spherical solute-rich phase of radius  $R$  surrounded by a spherical shell of solute-poor phase between the radii  $R$  and  $L/2$ ; neither the lattice idealization nor its spherical approximation are not expected to affect the result of the scaling calculation. We respectively denote the “charge densities” of the solute-rich and solute-poor phases by  $\rho_r$  and  $\rho_p$ , and note that according to the definition of these “charge densities”,  $\rho_r < 0$  and  $\rho_p > 0$ . These densities are constant away from the interface and are independent of the period, because the coexisting concentrations of solutes and the production rate are both constants that are independent of the period. Since the width of the interface (of order of the nano-scale, molecular size [14]) is typically negligible compared with the mesoscale domain size, the interface region can be neglected in the estimation of the “electrostatic interaction”. As explained before, Eqs. 3 and 4 imply that the integrals of the “charge densities” over the entire volume of the cell must vanish. Due to periodicity, the net charge in each of the spherical regions of radius  $L/2$  must also vanish. This “charge neutrality” relates the charge densities  $\rho_r$  and  $\rho_p$ , and the radius of the solute-rich phase,  $R$ , to the period  $L$  by the following equation:

$$\frac{4\pi}{3}\rho_r R^3 + \frac{4\pi}{3}\rho_p \left( \left( \frac{L}{2} \right)^3 - R^3 \right) = 0 \Rightarrow R = \left( \frac{\rho_p}{\rho_p - \rho_r} \right)^{\frac{1}{3}} \frac{L}{2} \quad (11)$$

To calculate the electrostatic energy within a spherical region of radius  $L/2$ , we find the electric field inside the spherical region using Gauss’s law. Because the electric field  $E(r)$  is spherically symmetric, it is related to the total charge  $Q(r)$  enclosed by a sphere of radius  $r$  whose center is the center of the spherical region, so that  $E(r) \sim Q(r)/r^2$ . The total charge  $Q(r)$  can be easily expressed by  $\rho_r$ ,  $\rho_p$ ,  $R$ ,  $L/2$ , and  $r$  as:

$$Q(r) = \begin{cases} \frac{4\pi}{3}r^3\rho_r & r < R \\ \frac{4\pi}{3}R^3\rho_r + \frac{4\pi}{3}\rho_p(r^3 - R^3) & R < r < \frac{L}{2} \end{cases} \quad (12)$$

The field  $E(r) \sim Q(r)/r^2$ , together with Eq. 12 for the total charge  $Q(r)$ , the formula for the electrostatic energy  $U \sim \int |E|^2 dV$ , and the relation between  $R$  and  $L$  that is given by Eq. 11, allow us to determine the scaling of the electrostatic energy  $U$  as a function of the period  $L$ , as  $U \sim L^5$ . We note that we must minimize the total effective free energy of the system that includes many such spherical regions. For fixed system volume, the number of such regions is inversely proportional to the volume of the region which depends on  $L$ . Thus, we minimize the ratio of the free energy per region and the volume of the region; namely, the

effective free energy density of the regions. The volume of a spherical region scales as  $L^3$ , which implies that the electrostatic energy density  $u$ , (including the ratio  $k_B T k_d / D$  which defines the strength of the effective coulomb interactions in Eq. 9), scales as  $u \sim k_B T k_d L^2 / D$ . In addition to the effective electrostatic energy, each spherical regions of radius  $L/2$  contributes an interfacial free energy due to interfacial tension  $\sigma$ . The interfacial area between the solute-rich and solute-poor domains within each of the spherical regions scales as  $\sim R^2$  which itself scales as  $\sim L^2$  according to Eq. 11. Therefore, the contribution of the interfacial energy to the free energy density of each spherical region is  $\sim \sigma L^2 / L^3 = \sigma / L$ . Minimization of the sum of the interfacial energy density and the electrostatic energy density predicts the scaling of the region size  $L$  (and also the domain size of the solute-rich phase) that minimizes the effective free energy  $\tilde{F}$ , with the interfacial tension  $\sigma$ , the degradation rate  $k_d$  and the mobility  $D$ :

$$\frac{d}{dL} \left( \frac{\sigma}{L} + \frac{k_B T k_d L^2}{D} \right) = 0 \Rightarrow L \sim \left( \frac{\sigma D}{k_B T k_d} \right)^{\frac{1}{3}} \quad (13)$$

This prediction for the domain size gives the same dependence of  $L$  on all the system parameters as the one-dimensional calculations and theoretical predictions of an independent, statistical-mechanical arguments in [6].

## 1.2 Theoretical derivation of the shape and orientation of the noise ellipse

In the subsection above, we considered the physics of LLPS in non-equilibrium ternary systems. Our main conclusion was that the non-equilibrium nature of biological cells modifies the spatial distribution and shapes of the phase separated domains, but, to a very good approximation, does not change the coexisting concentrations within them, which are well-described by equilibrium thermodynamics. Therefore, the binodal and tie-lines of equilibrium phase diagrams predict the coexisting concentrations of a non-equilibrium system with overall (system averaged) solute concentrations of  $\bar{\phi}$  and  $\bar{\psi}$ , where the bar denotes an average over the system volume. The total volumes of each of the two phases, can then be determined from a set of linear equations that relates the overall solute concentrations to the volumes of the coexisting phases and the concentrations of solutes within them. It is important to note that the  $\bar{\cdot}$  sign denotes spatial average, so that  $\bar{\phi}(t) = (\int \phi(\vec{r}, t) d\vec{r}) / V$  and  $\bar{\psi}(t) = (\int \psi(\vec{r}, t) d\vec{r}) / V$  ( $V$  being the volume of the system) are in general time dependent. The spatial averaging should not be confused with averaging over time or microstates, which we denote here by  $\langle \cdot \rangle$ .

In this section, we focus on the time evolution of the overall concentrations  $\bar{\phi}(t)$  and  $\bar{\psi}(t)$  (described by Eqs. 3 and 4) that is due to temporal changes in the production of solute molecules by gene expression. Protein production involves transcription of mRNA molecules from genes followed by translation of these mRNA molecules to proteins, and their eventual degradation. Both the production and degradation terms represent biochemical processes that are stochastic, which means that the rates in the right hand side of Eqs. 3 and 4 fluctuate in time. In cells that are in steady state (e.g. non-dividing cells in a confluent layer), the mean values of the production and degradation rates are constant in time, and only the stochastic fluctuations are expected to lead to variations of  $\bar{\phi}$  and  $\bar{\psi}$ . Therefore, repeated measurements of the overall concentrations  $\bar{\phi}, \bar{\psi}$  in different cells or in a single cell at sufficiently long time intervals (longer than the decorrelation time of the fluctuations), result in a distribution of  $(\bar{\phi}, \bar{\psi})$  that is centered around the mean values  $(\langle \bar{\phi} \rangle, \langle \bar{\psi} \rangle)$ . This distribution can be visualized (and can be measured) as a scatter-plot in a  $\phi - \psi$  diagram that is superimposed on top of the phase diagram. Equivalently, the distribution defines an "envelope" that delimits a closed curve (in the 2d space of  $\phi$  and  $\psi$ ) of constant probability density that is  $1/e^2$  of the peak value of the probability density, which is maximal at the mean values  $(\langle \bar{\phi} \rangle, \langle \bar{\psi} \rangle)$ . As we explain in the main text, if the  $(\bar{\phi}, \bar{\psi})$  distribution is not circularly-symmetric, it has a direction in the  $\phi - \psi$  diagram in which it is most spread (i.e. noise is maximal). When this direction is aligned with the direction of the tie-lines of the phase diagram, concentration buffering is most effective (see Fig. 2 in the main text).

Here, we determine the shape of the distribution in the limit where only the production, and not the degradation, is noisy, and that the fluctuations of the different stochastic degrees of freedom are uncorrelated in time and small compared to their mean values, so that one obtains a Gaussian distribution of the concentrations  $(\bar{\phi}, \bar{\psi})$ . The  $1/e^2$  locus of this distribution is in general a tilted ellipse in the space of concentrations

$(\phi, \psi)$ . We use linear response theory to predict the lengths of the axes of the ellipse and its orientation. At the end of this subsection, we relate the limit we have considered to the biological mechanisms that generate the noise, primarily transcriptional bursting [4]. We argue that although a more detailed noise structure, inspired by transcriptional bursts, may modify the shape of the distribution of  $(\bar{\phi}, \bar{\psi})$ , it is nevertheless expected to be non-circularly symmetric. Therefore, the conclusion of our paper regarding the alignment of the  $(\bar{\phi}, \bar{\psi})$  distribution with the tie-lines will still be valid. We begin with a brief review of an important distinction between two types of biological noise: intrinsic and extrinsic.

The stochasticity in production of biological molecules can have many molecular origins. However, it is customary to classify the sources of the noise as either, "intrinsic" or "extrinsic" [15]. To understand the distinction between them, we consider a hypothetical scenario, in which at a given time  $t$ , two cells are exact replicas of each other: they are genetically identical, have the same spatial concentrations profiles of all biomolecules, etc. Despite this equivalence, the production rates of different molecules in these two cells are not identical, because processes that underlie the gene expression responsible to production of all proteins and RNA molecules in the cell, are inherently stochastic. Because of this stochasticity, the production rates of a pair of identical genes, in the two hypothetical, identical cells, will differ. This type of noise, which is inherent to the expression of a gene, is referred to as "intrinsic noise", and is uncorrelated with the intrinsic noise of the expression of any other gene or allele.

Such intrinsic noise, however, is not the only noise that affects the production rates of proteins. The scenario in which the cellular environment is identical for two cells is, as mentioned, completely hypothetical. This is because the cellular environment itself is the consequence of expression of many genes, each subject to their own intrinsic noise. Therefore, the concentrations of many molecules that can affect the rates of production of molecules in the cells, such as transcription factors and ribosomal subunits, fluctuate and contribute to overall production noise. These environmental factors are "external" to the gene expression process itself and are thus termed "extrinsic variables" (e.g., concentrations of transcription factors) that contribute "extrinsic noise". Unlike the intrinsic noise of the expression of different genes, which is uncorrelated by definition, the extrinsic noise of the expression of different genes can be correlated if it originates from the same extrinsic variables.

To account for the intrinsic noise, we write the production rates of solutes  $\phi$  and  $\psi$  in Eqs. 3 and 4 as  $\bar{\Gamma}_\phi = \langle \bar{\Gamma}_\phi \rangle_i + \eta_\phi$  and  $\bar{\Gamma}_\psi = \langle \bar{\Gamma}_\psi \rangle_i + \eta_\psi$ , where  $\eta_\phi(t)$  and  $\eta_\psi(t)$  are stochastic terms that represent the intrinsic noise, and  $\langle \rangle_i$  denotes averaging over the intrinsic noise, but not the extrinsic one. For noise that is only delta-correlated in time [16], the stochastic terms of the intrinsic noise satisfy  $\langle \eta_\phi \rangle = \langle \eta_\psi \rangle = 0$ ,  $\langle \eta_\phi(t) \eta_\phi(t') \rangle = Q_\phi \delta(t - t')$ ,  $\langle \eta_\psi(t) \eta_\psi(t') \rangle = Q_\psi \delta(t - t')$ , and  $\langle \eta_\phi(t) \eta_\psi(t') \rangle = 0$ , where  $Q_\phi$  and  $Q_\psi$  are the magnitudes of the intrinsic noise terms, and  $\delta(t - t')$  is Dirac's delta function.

Due to the extrinsic noise, the production terms after averaging over the intrinsic noise,  $\langle \bar{\Gamma}_\phi \rangle_i$  and  $\langle \bar{\Gamma}_\psi \rangle_i$ , are still stochastic. To account for this additional stochasticity, we write the terms  $\langle \bar{\Gamma}_\phi \rangle_i$  and  $\langle \bar{\Gamma}_\psi \rangle_i$  as functions of the stochastic deviations from the mean value of the extrinsic variables,  $\langle \bar{\Gamma}_\phi(\eta_e(t)) \rangle_i$  and  $\langle \bar{\Gamma}_\psi(\eta_e(t)) \rangle_i$ . For notational simplicity, we consider only one such extrinsic variable denoted by  $\eta_e$ ; at the end of the subsection we explain how the results are modified for multiple extrinsic variables. We consider cells that are in steady state (e.g. non-dividing cells in confluent layers that are commonly used in experiments), so that the production rates  $\bar{\Gamma}_\phi$  and  $\bar{\Gamma}_\psi$  are only implicitly time-dependent via the stochastic variables  $\eta_e(t)$ ,  $\eta_\phi(t)$ , and  $\eta_\psi(t)$ . Because  $\eta_e(t)$  is a deviation from a mean value,  $\langle \eta_e(t) \rangle = 0$ . Furthermore, since  $\eta_e(t)$  represents intrinsic noise that affects the extrinsic variable,  $\eta_e(t)$  is uncorrelated with  $\eta_\phi(t)$ , and  $\eta_\psi(t)$ , so that  $\langle \eta_e(t) \eta_i(t') \rangle = 0$  and  $\langle \eta_e(t) \eta_e(t') \rangle = Q_e \delta(t - t')$ , where  $i = \phi, \psi$ ,  $Q_e$  is the magnitude of the extrinsic noise, and  $\delta(t - t')$  is Dirac's delta function. In general,  $\langle \bar{\Gamma}_\phi(\eta_e(t)) \rangle_i$  and  $\langle \bar{\Gamma}_\psi(\eta_e(t)) \rangle_i$  are non-linear functions of the extrinsic variables. However, using linear response theory we can expand these functions for a small deviation  $\eta_e(t)$  from the mean values of the extrinsic variable,  $\langle \bar{\Gamma}_\phi(\eta_e(t)) \rangle_i = \langle \bar{\Gamma}_\phi \rangle + R_\phi \eta_e(t)$  and  $\langle \bar{\Gamma}_\psi(\eta_e(t)) \rangle_i = \langle \bar{\Gamma}_\psi \rangle + R_\psi \eta_e(t)$ ; here,  $\langle \rangle$  denotes an average over all stochastic degrees of freedom so that  $\langle \bar{\Gamma}_\phi \rangle$  and  $\langle \bar{\Gamma}_\psi \rangle$  are no longer stochastic and, for cells in steady state, are time-independent.  $R_i = (\partial \bar{\Gamma}_i / \partial \eta_e) |_{\eta_e=0}$  is the linear response coefficient (which we consider to be constant to lowest order) of the production of solute  $i = \phi, \psi$  for small changes in the extrinsic variable. Substituting these expansions of the production terms into Eqs. 3 and 4 results in the following pair of stochastic equations for the production of the proteins  $\phi$

and  $\psi$ , which are related by their common source of extrinsic noise:

$$\frac{d\bar{\phi}(t)}{dt} = \langle \bar{\Gamma}_\phi \rangle - k_d^\phi \bar{\phi}(t) + R_\phi \eta_e(t) + \eta_\phi(t) \quad (14)$$

$$\frac{d\bar{\psi}(t)}{dt} = \langle \bar{\Gamma}_\psi \rangle - k_d^\psi \bar{\psi}(t) + R_\psi \eta_e(t) + \eta_\psi(t) \quad (15)$$

The mean, steady-state solution of Eqs. 14 and 15 gives  $\langle \bar{\phi} \rangle = \langle \bar{\Gamma}_\phi \rangle / k_d^\phi$  and  $\langle \bar{\psi} \rangle = \langle \bar{\Gamma}_\psi \rangle / k_d^\psi$ , which we respectively denote, as in the previous subsection, by  $\langle \bar{\Phi} \rangle$  and  $\langle \bar{\Psi} \rangle$ . This motivates us to respectively define  $\delta\bar{\phi}(t) = \bar{\phi}(t) - \langle \bar{\Phi} \rangle$  and  $\delta\bar{\psi}(t) = \bar{\psi}(t) - \langle \bar{\Psi} \rangle$  as the deviations of  $\bar{\phi}$  and  $\bar{\psi}$  from their mean, steady-state values. Using the definitions of  $\delta\bar{\phi}(t)$  and  $\delta\bar{\psi}(t)$ , as well as the following notation for the sums of the noise terms,  $N_\phi = R_\phi \eta_e(t) + \eta_\phi(t)$  and  $N_\psi = R_\psi \eta_e(t) + \eta_\psi(t)$ , we rewrite Eqs. 14 and 15 as two Langevin equations that are correlated by their noise terms:

$$\frac{d\delta\bar{\phi}(t)}{dt} = -k_d^\phi \delta\bar{\phi}(t) + N_\phi(t) \quad (16)$$

$$\frac{d\delta\bar{\psi}(t)}{dt} = -k_d^\psi \delta\bar{\psi}(t) + N_\psi(t) \quad (17)$$

The correlations and moments of the total noise terms  $N_\phi$  and  $N_\psi$ , which are determined by those of the separate, intrinsic and extrinsic noise terms, are written in the following form:  $\langle N_i(t) \rangle = 0$  and  $\langle N_i(t) N_j(t') \rangle = (Q_i \delta_{ij} + R_i R_j Q_e) \delta(t - t')$ , where  $i, j = \phi, \psi$ ,  $\delta_{ij}$  is Kronecker's delta, and  $\delta(t - t')$  is Dirac's delta function.

Using a linear transformation of the variables  $\delta\bar{\phi}(t)$  and  $\delta\bar{\psi}(t)$  that appear in Eqs. 16 and 17, it is possible to obtain a set of two linear, first-order, coupled equations of the new variables (i.e. such that both variables appear in each of the two equations), whose two noise terms are uncorrelated. These two equivalent and transformed Langevin equations represent an Ornstein-Uhlenbeck process with two variables (chapter 3 of [16]). Due to the linearity of the equations for an Ornstein-Uhlenbeck process, the steady-state, joint probability to measure specific values of both of its variables is a generalized Gaussian distribution (chapter 6 of [16]). Because the variables of the Ornstein-Uhlenbeck process are related to  $\delta\bar{\phi}(t)$  and  $\delta\bar{\psi}(t)$  by a linear transformation, the joint probability distribution for measurements of specific values of  $\delta\bar{\phi}(t)$  and  $\delta\bar{\psi}(t)$ ,  $P(\delta\bar{\phi}(t), \delta\bar{\psi}(t))$ , is also a generalized Gaussian of these two variables.

$$P(\delta\bar{\phi}, \delta\bar{\psi}) = \frac{1}{Z} \exp(-a_{\phi\phi} \delta\bar{\phi}^2 - a_{\phi\psi} \delta\bar{\phi} \delta\bar{\psi} - a_{\psi\psi} \delta\bar{\psi}^2) \quad (18)$$

where  $Z$  is a normalization constant, and  $a_{\phi\phi}$ ,  $a_{\phi\psi}$ , and  $a_{\psi\psi}$  are coefficients that are related to the  $\delta\bar{\phi}(t) - \delta\bar{\phi}(t)$ ,  $\delta\bar{\phi}(t) - \delta\bar{\psi}(t)$ , and  $\delta\bar{\psi}(t) - \delta\bar{\psi}(t)$  correlations (their variances and covariance), which can be found from the solutions of Eqs. 16 and 17. Eqs. 16 and 17, are linear, first-order, non-homogeneous, ordinary differential equations, thus their solutions for any specific realizations of the stochastic noise terms,  $N_\phi(t)$  and  $N_\psi(t)$ , are given by  $\delta\bar{\phi}(t) = \int_{-\infty}^t N_\phi(t') e^{k_d^\phi(t'-t)} dt'$  and  $\delta\bar{\psi}(t) = \int_{-\infty}^t N_\psi(t') e^{k_d^\psi(t'-t)} dt'$  [17]. These solutions, along with the statistical properties of  $N_\phi(t)$  and  $N_\psi(t)$ , suggest that the different correlations of  $\delta\bar{\phi}(t)$  and  $\delta\bar{\psi}(t)$  are  $\langle \delta\bar{\phi}(t)^2 \rangle = (Q_\phi + R_\phi^2 Q_e) / 2k_d^\phi$ ,  $\langle \delta\bar{\psi}(t)^2 \rangle = (Q_\psi + R_\psi^2 Q_e) / 2k_d^\psi$ ,  $\langle \delta\bar{\phi}(t) \delta\bar{\psi}(t) \rangle = R_\phi R_\psi Q_e / (k_d^\phi + k_d^\psi)$ . In turn, these values of the correlations imply that the values of the coefficients  $a_{\phi\phi}$ ,  $a_{\phi\psi}$ , and  $a_{\psi\psi}$  in the joint probability distribution of Eq. 18 are the following:

$$a_{\phi\phi} = \frac{k_d^\phi (k_d^\psi + k_d^\phi)^2 (Q_\psi + R_\psi^2 Q_e)}{(k_d^\psi + k_d^\phi)^2 (Q_\phi + R_\phi^2 Q_e) (Q_\psi + R_\psi^2 Q_e) - 4k_d^\phi k_d^\psi (R_\phi R_\psi Q_e)^2} \quad (19)$$

$$a_{\phi\psi} = -\frac{4k_d^\phi k_d^\psi (k_d^\psi + k_d^\phi) R_\phi R_\psi Q_e}{(k_d^\psi + k_d^\phi)^2 (Q_\phi + R_\phi^2 Q_e) (Q_\psi + R_\psi^2 Q_e) - 4k_d^\phi k_d^\psi (R_\phi R_\psi Q_e)^2} \quad (20)$$



$$a_{\psi\psi} = \frac{k_d^\psi (k_d^\psi + k_d^\phi)^2 (Q_\phi + R_\phi^2 Q_e)}{(k_d^\psi + k_d^\phi)^2 (Q_\phi + R_\phi^2 Q_e) (Q_\psi + R_\psi^2 Q_e) - 4k_d^\phi k_d^\psi (R_\phi R_\psi Q_e)^2} \quad (21)$$

In the case that multiple extrinsic variables (denoted by the index  $\alpha$ ) are considered, the results presented for the three coefficients above are modified by the following replacements  $R_\phi^2 Q_e \rightarrow \sum R_{\phi,\alpha}^2 Q_\alpha$ ,  $R_\psi^2 Q_e \rightarrow \sum R_{\psi,\alpha}^2 Q_\alpha$ , and  $R_\phi R_\psi Q_e \rightarrow \sum R_{\phi,\alpha} R_{\psi,\alpha} Q_\alpha$ ; the summations are over the different extrinsic noise variables  $\alpha$ , where the magnitude of the delta-correlated noise associated with each extrinsic variable is  $Q_\alpha$ , and the respective linear response coefficients for the productions of solutes  $\phi$  and  $\psi$ , for small changes in the extrinsic variable, are  $R_{\phi,\alpha}$  and  $R_{\psi,\alpha}$ .

The noise ellipse is derived from the distribution function of Eq. 18 by finding the locus of points in the  $\phi - \psi$  concentration diagram for which the distribution function has a value of  $1/e^2$  of its maximum (for  $\delta\bar{\phi} = \delta\bar{\psi} = 0$ ). This locus is determined from the equation  $a_{\phi\phi}\delta\bar{\phi}^2 + a_{\phi\psi}\delta\bar{\phi}\delta\bar{\psi} + a_{\psi\psi}\delta\bar{\psi}^2 = 2$ , Which indeed represents a general ellipse in the  $\phi - \psi$  diagram. The respective lengths of the major and minor axes,  $D_{ma}$  and  $D_{mi}$ , and the tilt angle  $\theta$ , are given in terms of the coefficients  $a_{\phi\phi}$ ,  $a_{\phi\psi}$ , and  $a_{\psi\psi}$  [18] by:

$$D_{ma} = 2 \sqrt{\frac{(a_{\phi\phi} + a_{\psi\psi}) + \sqrt{(a_{\phi\phi} + a_{\psi\psi})^2 - (4a_{\phi\phi}a_{\psi\psi} - a_{\phi\psi}^2)}}{(4a_{\phi\phi}a_{\psi\psi} - a_{\phi\psi}^2)}} \quad (22)$$

$$D_{mi} = 2 \sqrt{\frac{(a_{\phi\phi} + a_{\psi\psi}) - \sqrt{(a_{\phi\phi} + a_{\psi\psi})^2 - (4a_{\phi\phi}a_{\psi\psi} - a_{\phi\psi}^2)}}{(4a_{\phi\phi}a_{\psi\psi} - a_{\phi\psi}^2)}} \quad (23)$$

$$\varphi = \begin{cases} \frac{1}{2} \operatorname{arccot}((a_{\phi\phi} - a_{\psi\psi})/a_{\phi\psi}) & \text{If } a_{\phi\psi} \neq 0 \text{ and } a_{\phi\phi} < a_{\psi\psi} \\ 90^\circ + \frac{1}{2} \operatorname{arccot}((a_{\phi\phi} - a_{\psi\psi})/a_{\phi\psi}) & \text{If } a_{\phi\psi} \neq 0 \text{ and } a_{\phi\phi} > a_{\psi\psi} \\ 0^\circ & \text{If } a_{\phi\psi} = 0 \text{ and } a_{\phi\phi} < a_{\psi\psi} \\ 90^\circ & \text{If } a_{\phi\psi} = 0 \text{ and } a_{\phi\phi} > a_{\psi\psi} \end{cases} \quad (24)$$

As a mathematical side-note, we comment that the validity of the results above requires that the coefficients satisfy  $4a_{\phi\phi}a_{\psi\psi} - a_{\phi\psi}^2 > 0$ ; this inequality is always satisfied by the values of  $a_{\phi\phi}$ ,  $a_{\phi\psi}$ , and  $a_{\psi\psi}$ , that are described by Eqs. 19, 20, and 21.

Now that we have derived the shape and orientation of the "noise-ellipse", we consider two specific cases of interest. In the first case, we consider the scenario that there is no extrinsic noise,  $Q_e = 0$  (or  $Q_\alpha = 0$  for any  $\alpha$  in the case of multiple extrinsic variables). For  $Q_e = 0$ , Eq. 20 indicates that  $a_{\phi\psi} = 0$  and consequently Eq. 24 predicts that the tilt angle of the noise ellipse is either  $0^\circ$  or  $90^\circ$ , depending on the relative magnitudes of the intrinsic noises of the expressions of  $\phi$  and  $\psi$ . This result is expected, since in the absence of extrinsic noise we expect the fluctuations in the concentrations of  $\phi$  and  $\psi$  to be uncorrelated, so that the concentration of each of the two solutes fluctuates independently, leading to a noise ellipse whose major and minor axes are parallel to the  $\phi$  and  $\psi$  axes.

In the second case, we consider, for simplicity, the scenario that  $\phi$  and  $\psi$  are distinguishable molecules but identical in their degradation rates and noise-generating mechanisms so that  $k_d^\psi = k_d^\phi \equiv k_d$ ,  $Q_\phi = Q_\psi \equiv Q$  and  $R_\phi = R_\psi \equiv R$  (in the multiple extrinsic noise variables,  $R_{\phi,\alpha} = R_{\psi,\alpha} \equiv R_\alpha$  for any  $\alpha$ ). This scenario corresponds to the popular two-reporter assay that is used for measurements of the extrinsic and intrinsic noises related to expression of a specific gene [19]. In that case,  $a_{\phi\phi} = a_{\psi\psi}$ , and the coefficients  $a_{\phi\phi}$  and  $a_{\phi\psi}$  can be written in a simplified manner:

$$a_{\phi\phi} = \frac{k_d (Q + R^2 Q_e)}{Q (Q + 2R^2 Q_e)} \quad a_{\phi\psi} = -\frac{2k_d R^2 Q_e}{Q (Q + 2R^2 Q_e)} \quad (25)$$

where the lengths of the major and minor axes of the noise ellipse are respectively:

$$D_{ma} = \sqrt{\frac{2(Q + 2R^2 Q_e)}{k_d}} \quad D_{mi} = \sqrt{\frac{2Q}{k_d}} \quad (26)$$

and the tilt angle of the noise ellipse is  $\varphi = 45^\circ$ . The squared length of the minor axis  $D_{mi}^2$  arises from the intrinsic noise alone, while the squared length of the major axis  $D_{ma}^2$  is related to sum of the contributions of both the intrinsic and extrinsic noises. These predictions are in complete agreement with those of the original work that suggested the two-reporter assay for distinguishing between intrinsic and extrinsic noises [15]. Our results generalize those of [15] by predicting the shape of the noise ellipse of a pair of proteins in the general case that their degradation rates, susceptibilities of their expressions to the extrinsic noise, or the magnitudes of their intrinsic noises are not identical.

To conclude this subsection, we highlight that the predictions for the shape and size of the noise ellipse derived here are valid under the following conditions: (1) All the stochastic noise terms are Gaussian and uncorrelated in time. (2) The fluctuations of the extrinsic variables are small compared to their mean values. (3) The stochastic fluctuations of the degradation rate are negligible compared to the stochastic fluctuations of the production rate.

Our choice in point (1) of Gaussian noise that is uncorrelated in time is often used in statistical dynamics [16] and allows us to obtain analytical expressions for the shape and orientation of the noise ellipse. However, we are aware that this noise structure is only an approximation to the biological noise, which is primarily the result of transcriptional bursts [4], arising from stochastic alternation of a gene between transcriptionally inactive and transcriptionally active states. On the average, transcriptional bursts occur once every several hours and give rise to  $\sim 100$  copies of mRNA [4], which are then translated in the ribosomes and eventually degrade. The stochastic fluctuations of the production rates due to transcriptional bursts were previously expressed analytically in terms of the characteristics of the bursts (frequency and number of mRNA molecules produced), and mRNA lifetime, and found to be correlated in time and generally non-Gaussian [4]. However, as long as measurements of the concentrations are conducted on different cells or in a single cell in intervals that are longer than the time over which the mRNA or proteins degrade, more complex correlations of the noise over time are not expected to change the shape of the measured noise ellipse. Furthermore, the non-Gaussian nature of the noise generated by transcriptional bursts affect the contribution of higher moments of the noise which may modify the shape of the ellipse; but their effect is small if the magnitude of the noise is small as described in point (2).

Points (2) and (3) allow us to ignore higher order moments of the noise terms and terms beyond linear response of the production rates. These simplifications are essential to derive the analytic form of the noise ellipse. Inclusion of the higher order terms in the expansion of the productions with respect to the fluctuating extrinsic variables could modify the shape of the joint probability distribution for the concentrations, especially if the magnitudes of the noise terms are large. However, it is not expected to change the shape of the distribution to a one that is circularly symmetric. Therefore, the main result of our work, that alignment between the tie-lines of the phase diagrams and the direction of the envelope of the joint concentration distribution improves the efficacy of concentration buffering and may be selected by evolution, should still be valid even if the envelope is not an ellipse but another similarly shaped, closed curve.

## 2 Phase separation of systems with multiple solutes

The equilibrium thermodynamics of a system of  $N$  solutes in a solvent, whose concentrations are characterized by the set  $\{\phi_1, \dots, \phi_N\}$ , is described by its free energy density function  $f(\phi_1, \dots, \phi_N)$ . In equilibrium, the number of coexisting phases and their concentrations are determined by minimization of the total free energy (the spatial integral of  $f$ ) subject to conservation of the number of molecules of each species (including the solvent) and (for an incompressible system), their molecular volumes. If the equilibrium state of the system is a single, mixed phase of all the solutes in the solvent, then the equilibrium concentrations of the solutes are constant (neglecting local and transient thermal fluctuations), and the mean-field total free energy  $F$  is related to the free energy density by  $F = Vf(\phi_1, \dots, \phi_N)$ , where  $V$  is the volume of the entire system. In contrast, if the equilibrium state comprises multiple phases, then the total free energy is the sum of the products of the free energy densities and the volumes of each of the phases. When the interactions between the solutes are short ranged, the concentrations of solutes are constant within the bulk of the phase, and only vary across the interfaces between phases that, except very close to the critical point, have a width on the order

of the molecular size. In this case, the total free energy can be approximated as  $F = \sum_{j=1}^k V_j f(\phi_{1j}, \dots, \phi_{Nj})$ , where  $\phi_{ij}$  is the concentration of solute  $i$  in phase  $j$ ,  $V_j$  is volume of phase  $j$ , and  $k$  is the number of phases that coexist in equilibrium; we do not include gradient terms in the free energy density  $f$ , since those only affect the spatial organization of the phase-separated domains (see subsection 1.1 “Theory of non-equilibrium LLPS”) and not the solute concentrations within them, which are constant a few molecular sizes away from the interface.

The equilibrium state of the system is the one that minimizes the total free energy of the system, subject to the constraints that the total number of each molecular species and (for an incompressible system), the overall volumes of all the molecules (including solvent) in the system are constant. To satisfy these conditions, the free energy is supplemented by terms related to the chemical potentials and osmotic pressure, which results in the grand potential  $G$ :

$$G = \sum_{j=1}^k V_j f(\phi_{1j}, \dots, \phi_{Nj}) + \sum_{i=1}^N \frac{\eta_i}{v_0} \left( V \bar{\phi}_i - \sum_{j=1}^k V_j \phi_{ij} \right) - \frac{\pi}{v_0} \left( V - \sum_{j=1}^k V_j \right) \quad (27)$$

where  $v_0$  is a molecular volume of the order of the volume of a segment of the polymers (both  $\phi$  and  $\psi$ ) that is of the length of one persistence length,  $\bar{\phi}_i$  is the system-averaged, overall concentration of solute  $i$ ,  $\eta_i$  is the chemical potential of solute  $i$ , and  $\pi$  is the osmotic pressure of the system. Physically, the chemical potential of a solute represents the change in the free energy due to addition or removal of one molecule of that solute, and the osmotic pressure is the free energy density change due to local changes of volume (expansion or compression). In equilibrium, the chemical potentials and osmotic pressures of the coexisting phases must be equal. Otherwise, the total free energy will not be minimal, since fluxes of solutes or solvent molecules between phases can decrease the total free energy of the system.

Mathematically and equivalently, the equality of chemical potentials  $\eta_i$  and osmotic pressure  $\pi$  result from the minimization of the grand potential of Eq. 27 with respect to all of its variables: the concentrations of solutes in each of the coexisting phases,  $\phi_{ij}$ , and the volumes of the different phases,  $V_j$ . The values of the chemical potentials and osmotic pressure are then determined by the conservation constraints on number of each species and their molecular volumes. This results in the following equations:

$$\frac{\partial f(\phi_{1j}, \dots, \phi_{Nj})}{\partial \phi_{ij}} - \frac{\eta_i}{v_0} = 0 \quad (28)$$

$$f(\phi_{1j}, \dots, \phi_{Nj}) v_0 - \sum_{i=1}^N \eta_i \phi_{ij} + \pi = 0 \quad (29)$$

$$V \bar{\phi}_i - \sum_{j=1}^k V_j \phi_{ij} = 0 \quad (30)$$

$$V - \sum_{j=1}^k V_j = 0 \quad (31)$$

Eq. 28 implies the equality of the chemical potentials of each solute in each coexisting phase in equilibrium. Eq. 29 implies the equality of the osmotic pressures in each of the coexisting phases in equilibrium. It is important to note that Eqs. 28 and 29 are sufficient to determine the phase diagram of a system, namely the number of phases and coexisting concentrations for different overall concentrations of solutes. However, if one wishes to determine the actual volumes of the coexisting phases, Eqs. 30 and 31 must be used as well. These two equations enforce the conditions that the total molecular volume of each solute and the solvent remain constant (for an incompressible system).

We note that Eqs. 28, 29, 30, and 31, depend on the number of coexisting phases  $k$ , which is, a-priori, unknown. Furthermore, the single-phase state with  $k = 1$ ,  $V_1 = V$ , and  $\phi_{i1} = \bar{\phi}_i$ , always and trivially solves this set of equations. This implies that a system which shows LLPS can, in general, exhibit several states

which solve the non-linear set of Eqs. 28, 29, 30, and 31. Mathematically, this is because this set of equations characterizes the local extrema of the total free energy of the system rather than its global minimum. In principle, to find the global minimum one must obtain all the possible solutions to this set of equations with different values for  $k$  that ranges from 1 to  $N+1$  (we show in the next section that  $k$  cannot exceed  $N+1$ ) and compare their total free energies. This might be a non-trivial computational challenge. However, the convex “hull” surface which represents the minimal, total free energy as a function of the overall concentrations, provides a simple, computational way to find the global minimum for every combination of overall solute concentrations; for detailed explanation of this method, the reader is referred to [20].

In the section below, we predict the nature of the coexisting phases for multicomponent systems at fixed temperature and interactions, by investigating the mathematical structure of these four equations. We compare the number of equations (that originate in the physics) in Eqs. 28 and 29, and the number of unknown concentrations. In some cases, these numbers are not equal which suggests that the concentrations of the coexisting phases in the multicomponent system are not all specified, leading to implicit relations among them for fixed values of the interactions and temperature. In a simple system of one solute that separates into two phases ( $N = 1$  and  $k = 2$ ), the two coexisting concentrations are completely specified at each value of the temperature and interaction strength. However, for the two-solute systems considered in the main text ( $N = 2$ ), that phase separate into two phases ( $k = 2$ ), the binodals give a continuous range of values of the two solute concentrations that can exist in the two coexisting phases. As we discuss below, there is also a region in the concentration plane where three phases can coexist ( $k = 3$ ) for some  $N = 2$  systems. We show that this equilibrium specifies, for each of the three coexisting phases, only one pair of (or in general, a discrete set of pairs) solute concentrations and one solvent concentration, and not a continuous range as is the case for the region within the binodal of the two phase-equilibria discussed in the main text. The latter arises from what we mean by implicit relations above: In some of the phase separating systems, the physical equations do not specify specific (or a discrete set) of coexisting concentrations, but an entire range. These implicit relations allow us to define the binodal and tie-lines and discuss their geometry for general,  $N$ -dimensional phase diagrams that characterize the concentrations of  $N$  solutes that separate into  $k$  phases. We limit our analysis below to the case where the system phase separate, namely  $k \geq 2$ .

When  $k = N+1$  (e.g., three-phase equilibrium for two solutes), we show below that all the concentrations of the coexisting phases are specified for each region in the phase diagram where this equilibrium exists (there may be a discrete number of such regions). In that case, any noise in the concentrations that lies in the region of the  $N+1$  coexisting phases will still give rise to only that single set of solute and solvent concentrations. This represents complete concentration buffering. However, if the noise lies outside of the regions of the  $k = N+1$  coexisting phases, the buffering may not be complete and one must generalize the situation analogous to the alignment of the noise-ellipses and tie-lines for the case of two-solute systems, as discussed in the main text.

### 3 Geometry of phase diagrams of multicomponent systems that phase separate into multiple phases

The systems we consider in the main text comprise two solutes and a solvent ( $N = 2$ ), and phase separate into two phases ( $k = 2$ ). However, LLPS in cellular environment is usually more complicated. Biomolecular condensates (BMCs) commonly consist more than two solutes, and many times the proteins involved in LLPS undergo post-translational-modifications, which further increase the number of different species of solutes in the BMC ( $N > 2$ ). Furthermore, some BMCs are formed by LLPS that results in multiple phases ( $k > 2$ ). For example, the nucleolus has a layered structure that consists three distinct layers [21], and its LLPS leads to coexistence within once BMC of three phases (the layers of the nucleolus) that all coexist with a dilute phase of the same proteins ( $k = 4$ ). Therefore, it is important to analyze phase separating systems with  $N, k > 2$  that are more complex than ternary systems which phase separate into two phases. In this section, we investigate the geometry of the phase diagrams, binodals, and tie-lines for the general case of multicomponent, phase separating systems, and relate this to concentration buffering in such systems. We begin by examining a few simple examples of systems with two or three solutes and then generalize

the discussion using mathematical arguments to specify the geometry of the phase diagrams, binodals, and tie-lines for more complex, multicomponent systems.

### 3.1 Simple cases of $N = 2, 3$ -component systems that phase separate into $k = 2, 3$ , or 4 coexisting phases

We begin by analyzing the situation of a ternary system ( $N = 2$ ), which phase separates into three ( $k = 3$ ) phases. The situation of  $k = 2$  is analyzed quantitatively in the main text for various types of interactions (homotypic and/or heterotypic). At fixed values of the interactions and temperature, the phase diagrams consist of binodal curves (often loops) that enclose a region in the two-dimensional plane of the solute concentrations. However, this is not the case for three-phase coexistence in these systems. In that case, the coexisting phases are specified by three pairs of concentrations  $(\phi, \psi)$ , which represent six unknown concentrations. These concentrations are related by the six equations for the equality of chemical potentials (Eq. 28), three for  $\phi$  and three for  $\psi$ , and three equations for the equality of osmotic pressures (Eq. 29), a total of nine equations. However, these equations introduce three additional unknowns: two chemical potentials,  $\eta_\phi$  and  $\eta_\psi$ , and an osmotic pressure  $\pi$ . Therefore, the number of unknowns (six concentrations, two chemical potentials, and one osmotic pressure), is equal to the number of equations (six from Eq. 28 and three from Eq. 29), so that we expect a discrete number of solutions (sometimes only one) for this system of equations. In that case, there are three vertices that define a triangular region of concentration space with three-phase coexistence. Note that each such region determines a single set of solute and solvent concentrations, in contrast to the situation of two-phase equilibrium in these two-solute systems, where there is a continuous line of concentrations that specify the two possible coexisting phases. We call those vertices the branches of a multinodal. In the general case, a branch of the multinodal is the locus of points in the phase diagram that characterizes possible solute concentrations for one coexisting phase; the multinodal of a system that phase separates into  $k$  phases has  $k$  branches. To further explore how the geometry of the multinodal changes with the number of solutes or phases, we next analyze systems with three solutes. For such systems, the phase phase diagram is three-dimensional since each phase is characterized by three solute concentrations.

In these systems ( $N = 3$ ) that phase separate into two phases ( $k = 2$ ), the number of unknown concentrations is six, three for each of the two phases. Furthermore, there are three unknown chemical potentials, one for each solute type, and one unknown osmotic pressure, as before. These ten unknowns are related by eight equations, two equations for equality of chemical potentials for each of the three solutes, and two equations for the equality of osmotic pressures. Therefore, there are a total of eight equations and ten unknowns. The excess of two unknowns compared to the equations implies that the two branches of the binodal (the number of branches of the multinodal is equal to the number of coexisting phases), have two unspecified concentrations, so that they are two-dimensional surfaces. If the system ( $N = 3$ ) phase separates into three phases ( $k = 3$ ), there are three additional unknowns: the solute concentrations within the additional coexisting phase, and four additional equations, one for each of the three chemical potentials and the single osmotic pressures. Therefore, the excess of unknowns over equations is now one, suggesting that the three branches of the multinodal are one-dimensional curves. Similarly, if the system ( $N = 3$ ) phase separates into four phases ( $k = 4$ ), the numbers of unknowns and equations become equal, so that the four branches of the multinodal consist of zero-dimensional points, similar to the three-phase equilibrium considered in the previous paragraph for ternary systems ( $N = 2$ ). However, now each four-phase region is associated with four concentrations (three solutes and one solvent), and not three as before.

In addition to the geometry of the multinodal, it is important to consider the geometry of the tie-lines that connect coexisting concentrations in the case that the system consists many solutes and phase separates into multiple phases. For such general system, we define “tie-surfaces” as the loci of points in the phase diagram that characterize all the possible combinations of overall (system averaged) solute concentrations, in systems that phase separate into the same set of coexisting phases; as we show below, “tie-surfaces” are lines only in the case of two-phase equilibria ( $k = 2$ ). The possible sets of coexisting phases are described by corresponding points of equal chemical potentials and osmotic pressures on different branches of the multinodal. Once these corresponding points of a specific set are determined, the possible overall solute concentrations in systems

that phase separate into this specific set of coexisting phases are determined by varying the volumes in Eq. 30, subject to the constraints of Eq. 31 and positive volumes. If the number of equilibrium phases is two ( $k = 2$ ), then the tie-surfaces connect pairs of corresponding points on the two different branches of the binodal, and are thus lines, which are straight because Eq. 30 is linear in the volumes of the phases. This is true regardless of the number of components  $N$ . For the case of three-phase equilibria, the tie-surfaces connect three corresponding points and are thus planar triangles, and similarly, for four-phase equilibria, the tie-surfaces are triangular pyramids. As in the case of two-phase equilibria, this is true regardless of the number of components in the system.

In conclusion, we have shown for a few specific examples that an increase of the number of equilibrium phases decreases the dimensionality of the branches of the multinodal but increases their number (which is equal to the number of equilibrium phases). In contrast, increasing the number of solutes increases the dimensionality of the branches of the multinodal, as well as that of the phase diagram itself. Furthermore, we have argued that the geometry of the generalized “tie-surfaces” is determined by the number of equilibrium phases and not by the number of components in the system. In the subsection below, we demonstrate that in the general case of an  $N$ -component system that phase separates into  $k$  phases, the multinodal branches are  $(N - k + 1)$ -dimensional hyper-surfaces, and the tie-surfaces are  $(k - 1)$ -dimensional simplexes. Additionally, we briefly discuss how the geometry of the tie-surfaces relates to buffering of expression noise in an  $N$ -solute system.

### 3.2 General case of an $N$ -component system that phase separates into $k$ -phases

To mathematically prove the concluding statements of the previous subsection, we count the number of equations and unknowns for a general system of  $N$ -solute that phase separates into  $k$ -phases: Eqs. 28 and 29 respectively provide  $kN$  and  $k$  different equations (for different values of  $i = 1, \dots, N$  and  $j = 1, \dots, k$ ), in terms of a total of  $kN + N + 1$  unknowns:  $kN$  concentrations,  $N$  chemical potentials and 1 osmotic pressure.  $N + 1$  of these unknowns, the chemical potentials and osmotic pressure, do not appear in Eqs. 30 and 31. Therefore, it is convenient to eliminate the chemical potentials and osmotic pressure from Eqs. 28 and 29. This is achieved by two mathematical steps. First, the dependence on  $\eta_i$  in Eq. 29 for each value of  $j$ , is eliminated by substituting the expression of  $\eta_i$  from Eq. 28 with the same value of  $j$  into Eq. 29. Second, we choose an arbitrary value of  $j$  that characterize a “reference” phase, and subtract Eqs. 28 and 29 for that value of  $j$  from Eqs. 28 and 29 for all the other values of  $j$ . It is important to note that the choice of the reference phase is arbitrary and lack any physical significance. For the arbitrary choice of  $j = k$  as the reference phase (any other choice is equally valid), this results in the following set of equations:

$$\frac{\partial f(\phi_{1j}, \dots, \phi_{Nj})}{\partial \phi_{ij}} = \frac{\partial f(\phi_{1k}, \dots, \phi_{Nk})}{\partial \phi_{ik}} \quad (32)$$

$$f(\phi_{1j}, \dots, \phi_{Nj}) - \sum_{i=1}^N \phi_{ij} \frac{\partial f(\phi_{1j}, \dots, \phi_{Nj})}{\partial \phi_{ij}} = f(\phi_{1k}, \dots, \phi_{Nk}) - \sum_{i=1}^N \phi_{ik} \frac{\partial f(\phi_{1k}, \dots, \phi_{Nk})}{\partial \phi_{ik}} \quad (33)$$

where  $i = 1, \dots, N$ , and  $j = 1, \dots, k - 1$ , because  $j = k$  was chosen as the arbitrary value of  $j$ . This set provides  $(k - 1)(N + 1)$  equations, and has  $kN$  unknown concentrations. Thus, there is an excess of  $kN - (k - 1)(N + 1) = N + 1 - k$  unknowns compared with the number of equations.

This general treatment allows us to reach the following conclusions related to the number of equilibrium phases  $k$ : (1)  $k$  cannot exceed the number of components ( $N$  solutes + 1 solvent), because then there would be more equations than concentrations so that the system of equations is overdetermined and cannot be generally solved. This is physically reasonable because  $N + 1$  is the maximum number of independent concentrations of the  $N + 1$  components; additional phases beyond  $N + 1$  can “be incorporated” within the  $N + 1$  phases whose free energy densities are the lowest. This would decrease the total free energy and only change the volumes of each of the  $N + 1$  coexisting phases of lowest free energy densities but not their solute concentrations. (2) For the special case that the number of phases equals the number of components,  $k = N + 1$ , the numbers of equations and unknowns in Eqs. 32 and 33 are equal, so that the coexisting concentrations are determined

by these equations alone. This represents the situation where the concentrations in each of the coexisting phases are an inherent property of the interactions of the system, and are independent of the overall (system averaged) concentrations,  $\bar{\phi}_i$ . In such cases, the concentration buffering is complete because systems made with different overall solute concentrations phase separate into phases that always have the same solute concentrations. A single solute in a solvent that can phase separate into two coexisting phases, so that  $k = 2$  and  $N = 1$ , is a system of this kind [14]. (3) If  $k < N + 1$ , we have more unknowns than equations so that additional equations are needed to uniquely determine the coexisting concentrations. These are available from Eqs. 30 and 31, which respectively include  $N$  and 1 equations; these  $N + 1$  additional equations then relate the coexisting concentrations to the  $k$  different volumes of the  $k$  coexisting phases as new unknowns. Therefore, the  $N + 1$  Eqs. 30 and 31 supplement the  $(k - 1)(N + 1)$  Eqs. 32 and 33 for a total of  $k(N + 1)$  equations. This is now equal to the number of unknown concentrations  $kN$  in the  $k$  coexisting phases, in the set Eqs. 32 and 33, and  $k$  unknown coexisting volumes in the set Eqs. 30 and 31, leading to a unique solution of all the coexisting concentrations and all the volumes of the phases.

Although the  $(k - 1)(N + 1)$  equations of Eqs. 32 and 33 cannot be solved uniquely since the number of unknown concentrations,  $kN$  is larger than the number of equations (when  $k - N - 1 < 0$ ), they nevertheless provide  $(k - 1)(N + 1)$  useful equations that relate the  $kN$  concentrations. For each arbitrary choice of a reference phase, the  $kN$  concentrations of the entire system is divided to  $N$  “reference” concentrations that characterize the reference phase and the remaining  $(k - 1)N$  concentrations (that we denote as “other concentrations”) within the rest of the phases. Since the number of equations provided by Eqs. 32 and 33,  $(k - 1)(N + 1)$ , is always larger than the number of the “other” concentrations,  $(k - 1)N$ . Thus, a subset of  $(k - 1)N$  equations out of the total of  $(k - 1)(N + 1)$  equations can be used to implicitly express the “other” concentrations in terms of the reference concentrations.

In that case, these “other” concentrations (as functions of the reference concentrations) can be substituted into the remaining  $(k - 1)(N + 1) - (k - 1)N = k - 1$  equations that were not used to relate the “other” and reference concentrations. This yields  $k - 1$  implicit expressions that relate the concentrations of any chosen reference phase, and are independent of the “other” concentrations or the overall concentrations. Since any phase can be chosen as the reference phase, a set of  $k - 1$  expressions exists for any of the coexisting phases. Even though these implicit expressions usually cannot be written as closed forms, their existence is important since they characterize the different coexisting branches of the multinodal in the phase diagrams of the multicomponent system. These  $k - 1$  expressions that relate the  $N$  solute concentrations within a coexisting phase, when plotted in an  $N$ -dimensional phase diagram, whose  $N$  axes are the  $N$  solute concentrations, characterize an  $(N - k + 1)$ -dimensional hyper-surface. The solute concentrations of each of the  $k$  phases are characterized by points on the hyper-surface of the phase. Thus, each of the  $(N - k + 1)$ -dimensional hyper-surfaces characterizes a different branch of the multinodal. There are a total of  $k(N - k + 1)$ -dimensional hyper-surfaces, one for each coexisting phase; together, they form the high-dimensional multinodal in the general case of a system of a solvent with  $N$  solutes that phase separates into  $k$  phases.

For any of the points on each of these  $(N - k + 1)$ -dimensional hyper-surfaces, there are  $k - 1$  corresponding points, one on each of the other branches of the multinodal. These sets of  $k$  points represent the concentrations of  $N$  solutes in  $k$  coexisting phases are related by equality of chemical potentials (Eq. 28), and osmotic pressures (Eq. 29). For every set of  $k$  coexisting points on the multinodal, it is useful to define the tie-surface as the locus of points in the  $N$ -dimensional diagram that represents overall solute concentrations of systems that phase separate into phases whose solute concentrations are described by these  $k$  points. Once a set of  $k$  corresponding points is determined from Eqs. 28 and 29, the tie-surfaces are can be plotted using Eqs. 30 and 31. This is done by varying the volumes of the phases in Eq. 30 under the physical constraint that the volumes are positive and satisfy Eq. 31. Because Eq. 30 is linear in the volumes of the phases, then in the general case of  $k$  phases, the tie-surfaces are  $(k - 1)$ -simplexes [22].

Furthermore, in the general case that the system has  $N$  solutes, the noise-ellipse becomes an  $N$ -dimensional noise-ellipsoid with  $N$  axes. However, since one-dimensional tie-lines characterize phase diagrams of any system that phase separates into two phases, then, even for this general case, there still is a single orientation of the tie-lines that, when aligned with the largest axis of the  $N$ -dimensional noise-ellipsoid, will result in the maximally effective concentration buffering. Therefore, our evolutionary hypothesis presented in the main text, can be simply generalized to this case. When the number of equilibrium phases,  $k$ , is larger

than two, the tie-lines become  $(k - 1)$ -simplexes, and thus have  $k - 1$  directions in concentration space that buffer concentration fluctuations. This suggests that in addition to the noise associated with the largest axis, noise associated with  $k - 2$  additional axes of the noise-ellipsoid can be buffered, and facilitate more effective concentration buffering. Similar to the evolutionary hypothesis presented in the main text for  $k = 2$ , mutations that change the orientations of these  $(k - 1)$ -simplexes in concentration space and result in their alignment with the  $k - 1$  largest axes of the noise-ellipse, may be evolutionary selected. In the extreme limit that  $k = N + 1$ , the tie-surfaces are  $N$ -dimensional volumes that in some cases, might fully enclose the noise-ellipsoid, which is also  $N$ -dimensional. In that situation, the noise may be completely buffered by changes of the volumes of the coexisting phases without any changes of the concentrations of solutes within them. In this case, the coexisting concentrations are independent of the overall concentrations, as indeed was shown above to be the case for  $k = N + 1$ .

## 4 Flory Huggins theory generalized to multiple solutes

So far, we have not specified the free energy density, so that our previous conclusions above are generic to any system of molecules with short-range interactions. However, to determine the phase diagrams for a specific system, an appropriate model for the free energy density must be specified. The focus of this paper is cellular systems. In such systems, the phase separating molecules are usually linear, intrinsically disordered proteins (IDPs) and RNA molecules. Their conformations are similar to those of flexible polymers (and are therefore fractal) so that the individual molecules are not compact and can interpenetrate [23]. In that case, monomers can interact equally well with each other whether they belong to the same polymer or to different ones. We note that this would not be the case for globular proteins which would only interact via their surface amino-acids and not interpenetrate. This suggests that a mean-field treatment of the interactions of polymers that depend only on the local concentrations of each of the different types of macromolecules in the system is a good approximation for the system. This is precisely the physical picture that underlies the mean-field Flory Huggins (FH) model of polymer mixtures, which uses a lattice model to describe the molecular microstates of the system [24]. The lattice size is typically taken to be of order the persistence length of the chains which we take to be identical. The solvent molecules, which are typically much smaller than the size of one persistence length of the chains, are treated as a “continuum” that fill any volume that is left after placing the two solutes on the lattice. Since, in the lattice model, the smallest volume that can be filled by the “continuum” solvent is one lattice site, the fundamental units of solvent considered in our theory are “blocks” containing many molecules that together occupy one lattice site. In the framework of the FH model, the free energy density for a multicomponent system of  $N$  solutes and a solvent is:

$$fv_0 = \sum_{i=1}^N \frac{\phi_i}{N_i} \log \left( \frac{\phi_i}{N_i} \right) + \phi_s \log(\phi_s) - \frac{1}{2} \sum_{i,j=1}^N \tilde{\chi}_{ij} \phi_i \phi_j - \sum_{i=1}^N \tilde{\chi}_{is} \phi_i \phi_s - \frac{1}{2} \tilde{\chi}_{ss} \phi_s^2 \quad (34)$$

where  $v_0$  is a molecular volume of the order of the volume of a segment of the polymers  $\phi_i$  that is of the length of one persistence length,  $\phi_i$  is the volume fraction of solute  $i$ ,  $N_i$  is the size of solute  $i$  in units of persistence lengths,  $\phi_s$  is the volume fraction of the solvent,  $\tilde{\chi}_{ij}$  is the interaction energy of monomers (persistence length segments of the size of one lattice constant) of solutes  $i$  and  $j$ ,  $\tilde{\chi}_{is}$  is the interaction energy of a “block” of solvent molecules and a monomer of solute  $i$ , and  $\tilde{\chi}_{ss}$  is the interaction energy of two solvent “blocks” [24]; positive values of  $\tilde{\chi}_{ij}$ ,  $\tilde{\chi}_{is}$ , and  $\tilde{\chi}_{ss}$ , represent attractive interactions and negative ones, repulsive interactions. Since the solutes and solvent are the only components in the system, their local volume fractions must sum to unity (within the incompressible lattice model). Therefore, the volume fraction of the solvent can be expressed as  $\phi_s = 1 - \sum_{i=1}^N \phi_i$ . Substitution of this expression into the free energy in Eq. 34, results in the following simplified expression for the FH free energy with the explicit dependence on the solvent volume fraction,  $\phi_s$ , eliminated:

$$fv_0 = \sum_{i=1}^N \frac{\phi_i}{N_i} \log(\phi_i) + \left( 1 - \sum_{i=1}^N \phi_i \right) \log \left( 1 - \sum_{i=1}^N \phi_i \right) - \frac{1}{2} \sum_{i,j=1}^N \chi_{ij} \phi_i \phi_j \quad (35)$$



where constant terms and terms which are linear in the  $\phi_i$ -s were omitted because they do not affect the LLPS thermodynamics. [Linear terms in the concentration merely renormalize the chemical potentials.]

In Eq. 35, the Flory interaction term  $\chi_{ij}$  describes an effective interaction between monomers of solutes  $i$  and  $j$ , relative to their interactions with the solvent,  $\chi_{ij} \equiv \tilde{\chi}_{ij} + \chi_{ss} - \chi_{is} - \chi_{js}$ . If  $i \neq j$ ,  $\chi_{ij}$  describes heterotypic interactions, while if  $i = j$ ,  $\chi_{ij}$  describes homotypic interactions. We note that, as captured by the model of stickers and spacers (see section ‘‘Theoretical model of multicomponent LLPS’’ in the main text), each phase-separating solute in biological cells can be thought of as a heterogeneous polymer (due to their different amino acid sequences), so that different monomers comprising a given solute are chemically distinct and may in principle, each interact with differing energies. However, in the mean-field spirit of the FH model, the interaction parameters  $\chi_{ij}$  are interpreted as interaction energies that are averaged over all the different pairs of monomers of solutes  $i$  and  $j$ .

Substituting the FH free energy into the generic equations for the binodal, Eqs. 28 and 29, results in a set of equations whose solution predicts the concentrations of each species in each of the coexisting phases. These are functions of the solute sizes and interactions:

$$\frac{1}{N_i} \log(\phi_{i\ell}) - \log\left(1 - \sum_{i=1}^N \phi_{i\ell}\right) - \sum_{j=1}^N \chi_{ij} \phi_{j\ell} + \frac{1}{N_i} - 1 = \eta_i \quad (36)$$

$$\log\left(1 - \sum_{i=1}^N \phi_{i\ell}\right) + \sum_{i=1}^N \left(1 - \frac{1}{N_i}\right) \phi_{i\ell} + \frac{1}{2} \sum_{i,j=1}^N \chi_{ij} \phi_{i\ell} \phi_{j\ell} = -\pi \quad (37)$$

where  $i$  and  $j$  are indices that denote types of solutes which range from 1 to  $N$ , and  $\ell$  is an index that denotes phases which ranges from 1 to the number of phases  $k$ . Once the concentrations in each of the coexisting phases are determined from Eqs. 36 and 37, the volumes of the coexisting concentrations are determined by Eqs. 30 and 31 as in the previous sections.

For the case of ternary systems of two solutes in a solvent that is the focus of this paper, Eq. 35 simplifies to yield the following form which we use in the main text:

$$f(\phi, \psi) v_0 = \frac{\phi}{N_\phi} \log(\phi) + \frac{\psi}{N_\psi} \log(\psi) + (1 - \phi - \psi) \log(1 - \phi - \psi) - \frac{1}{2} \chi_{\phi\phi} \phi^2 - \chi_{\phi\psi} \phi\psi - \frac{1}{2} \chi_{\psi\psi} \psi^2 \quad (38)$$

where  $\phi_1, \phi_2, N_1$ , and  $N_2$ , are respectively replaced by the simpler notations of  $\phi, \psi, N_\phi$ , and  $N_\psi$ . For the common case of two-phase equilibria, which characterizes most BMCs, the equations for the coexisting concentrations, Eqs. 36 and 37 become:

$$\frac{1}{N_\phi} \log\left(\frac{\phi_1}{\phi_2}\right) - \log\left(\frac{1 - \phi_1 - \psi_1}{1 - \phi_2 - \psi_2}\right) - \chi_{\phi\phi}(\phi_1 - \phi_2) - \chi_{\phi\psi}(\psi_1 - \psi_2) = 0 \quad (39)$$

$$\frac{1}{N_\psi} \log\left(\frac{\psi_1}{\psi_2}\right) - \log\left(\frac{1 - \phi_1 - \psi_1}{1 - \phi_2 - \psi_2}\right) - \chi_{\phi\psi}(\phi_1 - \phi_2) - \chi_{\psi\psi}(\psi_1 - \psi_2) = 0 \quad (40)$$

$$\log\left(\frac{1 - \phi_1 - \psi_1}{1 - \phi_2 - \psi_2}\right) + \left(1 - \frac{1}{N_\phi}\right)(\phi_1 - \phi_2) + \left(1 - \frac{1}{N_\psi}\right)(\psi_1 - \psi_2) + \quad (41)$$

$$\frac{1}{2} \chi_{\phi\phi}(\phi_1^2 - \phi_2^2) + \chi_{\phi\psi}(\phi_1\psi_1 - \phi_2\psi_2) + \frac{1}{2} \chi_{\psi\psi}(\psi_1^2 - \psi_2^2) = 0 \quad (42)$$

where the subscripts 1, 2 respectively label the two coexisting phases. This set of equations characterizes the binodal of the system. In order to find the volumes of the coexisting phases for a specific phase-separating system with overall solutes concentrations  $\bar{\phi}$  and  $\bar{\psi}$ , the equations for the conservation of solutes and volume, presented in Eqs. 30 and 31 for the general case, are used to supplement Eqs. 39. In the case of two-phase equilibria, these conservation equations obtain the following form:

$$V_1\phi_1 + V_2\phi_2 = V\bar{\phi} \quad (43)$$

$$V_1\psi_1 + V_2\psi_2 = V\bar{\psi} \quad (44)$$

$$V_1 + V_2 = V \quad (45)$$

## 5 Supplementary results

Motivated by the fact that many BMCs in nature contain no more than a single multivalent scaffold, we limit our discussion to ternary systems where the second solute does not have homotypic interactions,  $\chi_{\psi\psi} = 0$ , so that only solute  $\phi$  may serve as a scaffold. As is apparent from the phase diagrams shown in the main text, ternary systems with  $\chi_{\psi\psi} = 0$  can phase separate into two phases, but not three or more. This is consistent with the biological observation that to date, most BMCs are a single, highly concentrated phase that coexists in equilibrium with a dilute phase. A minority of BMCs, such as the complete nucleolus [21], comprise multiple phases, and are outside of the scope of the quantitative analysis presented in this paper. However, as explained in subsection “General case of an  $N$ -component system that phase separates into  $k$ -phases” above, multi-phase equilibria may be modulated by evolutionary pressures to provide concentration buffering that is even more robust than two-phase equilibria. In this section, we theoretically analyze the different possible scenarios of LLPS in ternary systems with  $\chi_{\psi\psi} = 0$ , and derive analytical expressions for the critical homotypic and heterotypic interactions mentioned in the main text, that are associated with different topologies of the phase diagrams. Tables 1 and 2 summarize the results of this section. Table 1 lists these critical interactions as functions and the molecular sizes,  $N_\phi$  and  $N_\psi$ , and for some critical heterotypic interactions also of the homotypic interaction  $\chi_{\phi\phi}$ . Table 2 relates these critical interactions to the various, predicted topologies of the phase-diagrams.

We use two physical insights related to the topologies of the phase diagrams to analytically derive the critical interaction parameters. First, the lines  $\phi = 0$ ,  $\psi = 0$ , and  $\phi + \psi = 1$  that enclose the physically relevant region of the phase diagram, where the volume fractions of the solutes and solvent molecules are all positive, all correspond to different types of binary systems. These are respectively: (i) a system that has no  $\phi$  solute, (ii) a system with no  $\psi$  solute, (iii) a system that is a blend of  $\phi$  and  $\psi$  solutes, but with no solvent. Therefore, if the binodal of the ternary system intersects any of these lines, the corresponding binary system undergoes LLPS even without the component whose volume fraction is zero. In that case, the effective interaction energy between the two components of the corresponding binary system must be larger than the critical one predicted by FH theory for binary systems [14].

Second, we note that a system can phase separate if and only if there is a region (termed the spinodal) in the phase diagram where the mixed, uniform state of the system ceases to be a minimum of the free energy density  $f$ . In that region, a uniform, mixed system becomes unstable and undergoes spinodal decomposition [25]. In other words, the existence of a binodal for LLPS implies the existence of a spinodal region which allows us to use the spinodal as an indicator of LLPS in the following subsections.

When phase separation exists, the generalization of the “common tangent” construction [14] that gives rise to binary LLPS in multicomponent systems, implies that the free energy surface as function of the concentrations, has a concave (spinodal) region, surrounded, on both sides by a region that is convex. The latter extends to small concentrations where the single-phase must be asymptotically stable due to the dominance of the entropy of mixing at low enough concentrations. The existence of such a spinodal region can also be proven mathematically using a simple *reductio ad absurdum* argument: If the mixed phase is a minimum of the free energy for all possible concentration, then the Hessian matrix  $H$  of  $f$ , defined by

$$H = v_0 \begin{pmatrix} \frac{\partial^2 f}{\partial \phi^2} & \frac{\partial^2 f}{\partial \phi \partial \psi} \\ \frac{\partial^2 f}{\partial \phi \partial \psi} & \frac{\partial^2 f}{\partial \psi^2} \end{pmatrix} \quad (46)$$

is positive definite in the entire, two-dimensional space delineated by the concentration axes that give rise to the phase diagram (“concentration space”). In that case, the free energy density  $f$  is everywhere convex. Therefore, the total free energy of a two-phase system is always larger than the total free energy of a one-phase system with the same overall concentrations. This because the average free energy density of the two-phase system is characterized by points along a straight line constructed between two points on the

surface described by  $f(\phi, \psi)$ , which is larger than  $f$  itself for the same solute concentrations if  $f$  is a convex function. Those two points would represent the coexisting phases of a phase separated system and in the presence of an intervening convex region, binary LLPS cannot be thermodynamically favorable. Conversely, if there is a concave region, the system can undergo spinodal decomposition and thus also phase separate where the two concentrations corresponding to the coexisting phases indeed have lower free energies than any of the single phases along the line connecting them. Therefore, when a phase-separated state of the system, is favorable, there must be a region within the area enclosed by the binodal, in which the single-phase state is unstable and  $H$ , defined by Eq. 46, has at least one negative eigenvalue. On the boundary of this inner, convex region, namely the spinodal surface,  $H$  has at least one vanishing eigenvalue, so that the determinant of  $H$  vanishes there.

This correspondence between the binodal and spinodal regions, allows us to use the more easily calculated spinodal as an indicator for the existence of a phase-separated (binodal) regions of the ternary systems. We do this below and analyze each of the three cases considered in the main text separately to derive analytical expressions for the critical values of the heterotypic and homotypic interactions that characterize the different cases and phase diagram topologies.

### 5.1 Case I - Homotypic ( $\chi_{\phi\phi} > 0, \chi_{\phi\psi} = 0$ ): Concentrated phase rich in one solute (segregative LLPS)

In this case, the interaction parameters of the free energy Eq. 38 obey  $\chi_{\phi\phi} > 0, \chi_{\phi\psi} = 0, \chi_{\psi\psi} = 0$ . As explained in the main text, ternary systems in this case comprising solvent, a homotypically attracting-solute and a crowder (which has excluded volume interactions with both the other solute and the solvent but no heterotypic or homotypic interactions), may exhibit three types of phase diagrams, depending on the magnitude of the homotypic interaction: (1) For strong homotypic interaction,  $\chi_{\phi\phi} > \chi_{\phi\phi}^c$  (given in Table 1), the homotypically-attracting solute can phase separate from the solvent in the absence of the crowder, so that the binodal in the phase diagram intersects the line  $\psi = 0$ , which characterizes the absence of crowder. (2) The homotypic interaction is strong enough to facilitate LLPS in the presence of a crowder but too weak to promote LLPS in its absence:  $\chi_{\phi\phi}^c > \chi_{\phi\phi}$ . However, the system can show LLPS in the presence of a crowder whose maximum concentration obeys  $\phi + \psi = 1$ , which applies for zero solvent; this constrains  $\chi_{\phi\phi} > \chi_{\phi\phi}^m$  (given in Table 1). In that case, the binodal does not intersect the line  $\psi = 0$ , which characterizes the absence of crowder, but does intersect the line  $\phi + \psi = 1$ , which characterizes the case where there are only the two solutes and no solvent. (3) The homotypic interaction is too weak to promote LLPS even if the solvent is at its maximal concentration (i.e., solvent is replaced by the crowder), so there is no binodal. This happens for  $\chi_{\phi\phi}^m > \chi_{\phi\phi}$ . Below, we derive analytical expressions for critical interaction parameters,  $\chi_{\phi\phi}^c$  and  $\chi_{\phi\phi}^m$ , as functions of the sizes of the solute and the crowder,  $N_\phi$  and  $N_\psi$ .

In order to find  $\chi_{\phi\phi}^c$ , we consider the limiting free energy,  $f_{\psi=0}^I(\phi)$ , of the binary system for the case of zero crowder,  $\psi = 0$ . Eq. 38 then reduces to:

$$f_{\psi=0}^I(\phi) v_0 = \frac{\phi}{N_\phi} \log(\phi) + (1 - \phi) \log(1 - \phi) - \frac{1}{2} \chi_{\phi\phi} \phi^2 \quad (47)$$

Since the critical parameter  $\chi_{\phi\phi}^c$  distinguishes homotypically-attracting solutes that can undergo LLPS in a binary system from those that cannot, it is equivalent to interaction energy at the critical point of LLPS of the binary system, characterized by the line  $\psi = 0$ . Therefore,  $\chi_{\phi\phi}^c$  solves the system of equation that results from equating the second and third derivatives of Eq. 47, with respect to  $\phi$ , to zero [14]:

$$0 = \frac{1}{N_\phi \phi_c} + \frac{1}{(1 - \phi_c)} - \chi_{\phi\phi}^c \quad (48)$$

$$0 = -\frac{1}{N_\phi \phi_c^2} + \frac{1}{(1 - \phi_c)^2} \quad (49)$$

where the value of  $\phi_c$  that solves these equation is the concentration of the solute  $\phi$  at the critical point of

the binary system of solute  $\phi$  and a solvent.  $\chi_{\phi\phi}^c$  that solves these equations is given by:

$$\chi_{\phi\phi}^c = \left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2 \quad (50)$$

Similarly, the homotypic interaction  $\chi_{\phi\phi}^m$  distinguishes between ternary systems where the binodal intersects the line  $\phi + \psi = 1$  and those where the binodal does not. The line  $\phi + \psi = 1$  characterizes a binary blend of homotypic solute and client with no solvent. We therefore analyze the free energy density,  $f_{\phi+\psi=1}^I(\phi)$ , of this binary system that is derived from Eq. 38 by substituting  $\psi = 1 - \phi$ :

$$f_{\phi+\psi=1}^I(\phi) v_0 = \frac{\phi}{N_\phi} \log(\phi) + \frac{(1-\phi)}{N_\psi} \log(1-\phi) - \frac{1}{2} \chi_{\phi\phi} \phi^2 \quad (51)$$

As indicated by the intersection of the binodal of the ternary system with the line  $\phi + \psi = 1$ , a binary blend of  $\phi$  and  $\psi$  undergoes LLPS only when the homotypic interaction is larger than the value for the critical interaction strength of the blend (with no solvent), given by equating the second and third derivatives of Eq. 51, with respect to  $\phi$ , to zero. We denote by  $\chi_{\phi\phi}^m$  this critical value derived from the solution for  $\chi_{\phi\phi}$  of the following two equations:

$$0 = \frac{1}{N_\phi \phi_m} + \frac{1}{N_\psi (1 - \phi_m)} - \chi_{\phi\phi}^m \quad (52)$$

$$0 = -\frac{1}{N_\phi \phi_m^2} + \frac{1}{N_\psi (1 - \phi_m)^2} \quad (53)$$

where  $\phi_m$  is the concentration of solute  $\phi$  at the critical point of the binary blend. The solution of this system of equations determines the value of  $\chi_{\phi\phi}^m$ :

$$\chi_{\phi\phi}^m = \left(\frac{1}{\sqrt{N_\phi}} + \frac{1}{\sqrt{N_\psi}}\right)^2 \quad (54)$$

We note that  $\chi_{\phi\phi}^m$  decreases with increasing  $N_\psi$ . This indicates that the effect of crowders on LLPS of a homotypically-attracting solute increases with the size of the crowder. Furthermore,  $\chi_{\phi\phi}^m$  attains two important limiting values when  $N_\psi \rightarrow \infty$  and  $N_\psi = 1$ . For  $N_\psi \rightarrow \infty$ , the critical interaction  $\chi_{\phi\phi}^m$  approaches  $N_\phi^{-1}$ . This limiting value of  $\chi_{\phi\phi}^m$  is a lower bound on the homotypic interaction of a solute  $\phi$  that phase separates in the presence of crowders. This corresponds to an interaction of  $k_B T$  (the units in which we measure  $\chi_{ij}$ ) per polymer, which is of size  $N_\phi$ . Smaller values of the interaction are not effective in generating LLPS since then the entropy per polymer (also of order  $k_B T$  per polymer) dominates and promotes mixing in a single, homogeneous phase.

In the other limit, where  $N_\psi = 1$ , the two critical homotypic interactions,  $\chi_{\phi\phi}^c$  and  $\chi_{\phi\phi}^m$ , which are respectively described by Eqs. 50 and 54, become equal. This special case where the minimal homotypic interaction energy required for solute  $\phi$  to phase separate in the presence of crowders,  $\chi_{\phi\phi}^m$ , is equal to its critical interaction energy required to phase separate in the absence of crowders,  $\chi_{\phi\phi}^c$ , suggests that the crowders have no effect, at least for  $N_\psi = 1$ . It is easily seen that this indeed is the case, when we write the free energy for case I in this limit, which results from Eq. 38 by substitution of  $\chi_{\psi\psi} = \chi_{\phi\psi} = 0$  and  $N_\psi = 1$ . This can be written in the form:

$$f(\phi, \psi) v_0 = \frac{\phi}{N_\phi} \log(\phi) + (1 - \phi) \log(1 - \phi) - \frac{\chi_{\phi\phi}}{2} \phi^2 \quad (55)$$

$$+ (1 - \phi) \left( \frac{\psi}{1 - \phi} \log\left(\frac{\psi}{1 - \phi}\right) + \left(1 - \frac{\psi}{1 - \phi}\right) \log\left(1 - \frac{\psi}{1 - \phi}\right) \right) \quad (56)$$

The first three terms of the right hand side of the equation are the same as the FH free energy of a binary system of solute  $\phi$  with an effective solvent of volume fraction  $1 - \phi$ . This effective solvent comprises

the aqueous phase and the  $\psi$  solute which give rise to a mixing entropy of these two components of the effective solvent given by the last term of Eq. 55. This rewriting is only appropriate for  $N_\psi = 1$ , but has important physical consequences. Since the driving force for LLPS is the homotypic interaction of  $\phi$ , only the contribution of the first three terms in Eq. 55 is decreased by LLPS. The contribution of the fourth term in Eq. 55 cannot decrease because that is relevant only to the "internal" properties of the effective solvent. As a matter of fact, if we measure the  $\psi$  solvent relative to the volume fraction of the effective solvent  $\psi' = \psi / (1 - \phi)$ , we see that the first three terms of the free energy are independent of the last one (which thus can be minimized independently). This demonstrates that indeed, when  $N_\psi = 1$ , the addition of crowders, which appear only in the fourth term of Eq. 55, has no effect on the phase diagrams of the homotypically-attracting solute  $\phi$ .

In order to investigate the effect of the crowder size  $N_\psi$  for general values of  $N_\psi$ , we consider the spinodal, which is easier to calculate than the full binodal. The Hessian,  $H_I$ , of the system is obtained by substitution of the free energy in 38 with  $\chi_{\psi\psi} = \chi_{\phi\psi} = 0$ , into Eq. 46:

$$H_I = \begin{pmatrix} \frac{1}{N_\phi\phi} + \frac{1}{(1-\phi-\psi)} - \chi_{\phi\phi} & \frac{1}{(1-\phi-\psi)} \\ \frac{1}{(1-\phi-\psi)} & \frac{1}{N_\psi\psi} + \frac{1}{(1-\phi-\psi)} \end{pmatrix}$$

The spinodal is given by  $\det(H_I) = 0$ , which results in a quadratic function of  $\phi$  whose two roots are expressed as follows:

$$\phi_{\pm} = \frac{1}{\chi_{\phi\phi}} \left( \frac{1}{2} \left( \Psi \pm \sqrt{\Psi^2 - \frac{4}{N_\phi}} \right) + \frac{1}{N_\phi} \right) \quad (57)$$

where  $\Psi = \chi_{\phi\phi}(N_\psi - 1)\psi + \chi_{\phi\phi} - 1 - \frac{1}{N_\phi}$  is a parameter related to the concentration of the crowder, its size, and the homotypic interaction strength of the  $\phi$  solute and its size, but not to the concentration of  $\phi$ . These two roots,  $\phi_+$  and  $\phi_-$ , characterize the two branches (large and small concentrations of  $\phi$  respectively) of the parabolic spinodal. We note that  $\phi_+$  and  $\phi_-$  depend on the length of the crowder,  $N_\psi$ , and its concentration,  $\psi$ , only implicitly through the parameter  $\Psi$ . Thus, we conclude that the effects of increasing of the length of the crowder and increasing its concentration are equivalent. Furthermore, our analysis leads to the following conclusions: (1) When  $N_\psi = 1$  the expression of  $\Psi$  becomes independent of  $\psi$ , implying that the phase diagram of the homotypically-attracting solute  $\phi$  is independent of the concentration of crowder  $\psi$ , in agreement with our conclusions presented above. (2)  $\phi_+$  and  $\phi_-$  are respectively monotonically increasing and decreasing function of  $\Psi$ , so that an increase of the crowder concentration or its length makes the spinodal wider, causing it to cover increasing area of the phase diagram. This indicates that both increasing the concentration of the crowder or its length promote LLPS. (3) In the limit  $\Psi \rightarrow \infty$ ,  $\phi_- \rightarrow (\chi_{\phi\phi}N_\phi)^{-1}$ , which suggests that if  $\chi_{\phi\phi}N_\phi < 1$ , the spinodal is entirely outside of the physical domain of concentration space limited by  $\phi, \psi \geq 0$  and  $\phi + \psi \leq 1$ . This means that the homotypic interactions are too weak to promote LLPS, in agreement with our discussion regarding the limiting value of  $\chi_{\phi\phi}^m$ .

## 5.2 Case II - Heterotypic ( $\chi_{\phi\phi} = 0, \chi_{\phi\psi} > 0$ ): Concentrated phase rich in both solutes (associative LLPS)

As described in the main text, the binodals for this case are closed loops in the  $\phi$ - $\psi$  phase diagram (at constant values of  $\chi_{\phi\psi}$ ), that shrink as the heterotypic interaction strength,  $\chi_{\phi\psi}$ , decreases. This suggests that there is a critical value of the heterotypic interaction parameter  $\chi_{\phi\psi}^c$ , below which the binodal disappears; namely, the ternary system does not phase separate. To analytically express  $\chi_{\phi\psi}^c$  as a function of the sizes of the two solutes,  $N_\phi$  and  $N_\psi$ , we consider the spinodal of the system, which is easier to calculate. At the heterotypic interaction value of  $\chi_{\phi\psi} = \chi_{\phi\psi}^c$ , the binodal shrinks to a point, so that, as explained above, the spinodal must shrink to a point as well. Therefore, equating the area of the spinodal to zero provides an equation whose solution is  $\chi_{\phi\psi}^c$ .

For the heterotypic case, the free energy density is related to Eq. 38 by setting  $\chi_{\phi\phi} = \chi_{\psi\psi} = 0$ . For calculation of the spinodal, we find the Hessian,  $H_{II}$ , which is obtained by substituting this free energy into

Eq. 46:

$$H_{II} = \begin{pmatrix} \frac{1}{N_\phi \phi} + \frac{1}{1-\phi-\psi} & \frac{1}{1-\phi-\psi} - \chi_{\phi\psi} \\ \frac{1}{1-\phi-\psi} - \chi_{\phi\psi} & \frac{1}{N_\psi \psi} + \frac{1}{1-\phi-\psi} \end{pmatrix} \quad (58)$$

The spinodal is the locus of points in the  $\phi$ - $\psi$  plane defined by  $\det(H_{II}) = 0$ . In terms of the concentrations of the two solutes, their sizes, and the heterotypic interaction parameter, we obtain the following implicit relation:

$$(1 + \psi(N_\psi - 1)) + ((N_\phi - 1) + N_\psi N_\phi (\chi_{\phi\psi} \psi - (\chi_{\phi\psi} - 2)) \chi_{\phi\psi} \psi) \phi + N_\psi N_\phi \chi_{\phi\psi}^2 \psi \phi^2 = 0 \quad (59)$$

The left hand side is a quadratic function of  $\phi$ . The discriminant  $\Delta_{II}$ , of this quadratic form is:

$$\Delta_{II} = ((N_\phi - 1) + N_\psi N_\phi (\chi_{\phi\psi} \psi - (\chi_{\phi\psi} - 2)) \chi_{\phi\psi} \psi)^2 - 4N_\psi N_\phi \chi_{\phi\psi}^2 \psi (1 + \psi(N_\psi - 1)) \quad (60)$$

When the concentration of the solute  $\psi$  is such that the discriminant is positive, there are two values of  $\phi$  that solve Eq. 59 for the given value  $\psi$ . If the range of values of  $\psi$  for which the discriminant is positive, are in the region of physical concentrations  $0 \leq \psi \leq 1$ , then this range of  $\psi$  results in a closed spinodal curve whose two halves are related to the two solutions for  $\phi$ . The minimal and maximal values of  $\psi$  along those curved loops are those where the discriminant goes to zero. This is expected, because the binodal is a closed loop (see Fig. 4 in the main text), so that the spinodal, which is always enclosed by the binodal, must be a closed loop as well. We now determine the endpoint values of  $\psi$  (where the discriminant goes to zero), which characterize a transition from positive to negative discriminant. At these endpoints the discriminant vanishes so that they must be roots of Eq. 60. Since Eq. 60 is a quartic polynomial in  $\psi$ , it has four roots that can be analytically expressed as:

$$\psi_1 = \frac{1}{2} \left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) - \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (61)$$

$$\psi_2 = \frac{1}{2} \left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) + \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (62)$$

$$\psi_3 = \frac{1}{2} \left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) - \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (63)$$

$$\psi_4 = \frac{1}{2} \left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) + \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (64)$$

where we used the Solve function of Wolfram Mathematica software to obtain these roots. The values of  $\phi$  that correspond to the points along the spinodal at which  $\psi$  is minimal or maximal are calculated for these four values of  $\psi$  using Eq. 59 are:

$$\phi_1 = \frac{1}{2} \frac{1}{\sqrt{N_\phi} + 1} \left( 1 + \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (65)$$

$$\phi_2 = \frac{1}{2} \frac{1}{\sqrt{N_\phi} + 1} \left( 1 - \sqrt{\left( 1 - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2} \right) \quad (66)$$

$$\phi_3 = -\frac{1}{2} \frac{1}{\sqrt{N_\phi} - 1} \left( 1 + \sqrt{\left(1 - \frac{2}{\chi_{\phi\psi}} \left(1 - \frac{1}{\sqrt{N_\phi}}\right)\right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left(1 - \frac{1}{\sqrt{N_\phi}}\right)^2} \right) \quad (67)$$

$$\phi_4 = -\frac{1}{2} \frac{1}{\sqrt{N_\phi} - 1} \left( 1 - \sqrt{\left(1 - \frac{2}{\chi_{\phi\psi}} \left(1 - \frac{1}{\sqrt{N_\phi}}\right)\right)^2 - \frac{4}{N_\psi \chi_{\phi\psi}^2} \left(1 - \frac{1}{\sqrt{N_\phi}}\right)^2} \right) \quad (68)$$

The first two of these four pairs  $(\phi_i, \psi_i)$ , represent the points along the spinodal that correspond to the maximal and minimal value of  $\psi$ , while we find below that the two other pairs represent unphysical values. Since, for  $\chi_{\phi\psi} > \chi_{\phi\psi}^c$  the spinodal exists and is contained within the physical region of the concentration space, the first two points satisfy the inequalities  $\phi, \psi \geq 0$  and  $\phi + \psi \leq 1$  for any value of  $\chi_{\phi\psi} > \chi_{\phi\psi}^c$ . We examine the last two points by considering their limiting values as  $\chi_{\phi\psi} \rightarrow \infty$ , which shows that  $\phi_3$  and  $\phi_4$  are negative and therefore are not physically relevant. In conclusion, the points  $(\phi_1, \psi_1)$  and  $(\phi_2, \psi_2)$  represent the points on the spinodal for which the values of  $\psi$  are maximal or minimal. When the heterotypic interaction has the critical value,  $\chi_{\phi\psi} = \chi_{\phi\psi}^c$ , the spinodal shrinks to a point so that the two points of maximal and minimal  $\psi$  become identical. Equating the expressions for these two points,  $(\phi_1, \psi_1) = (\phi_2, \psi_2)$  to find when their difference vanishes, yields an equation for the critical value of  $\chi_{\phi\psi} = \chi_{\phi\psi}^c$  for which the spinodal vanishes, whose solution is:

$$\chi_{\phi\psi}^c = 2 \left( 1 + \frac{1}{\sqrt{N_\psi}} \right) \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \quad (69)$$

### 5.3 Case III - Homotypic and heterotypic interactions ( $\chi_{\phi\phi} > 0, \chi_{\phi\psi} > 0$ ): LLPS of a scaffold with a multivalent binding partner

As explained in the main text, ternary systems with a combination of both homotypic,  $\chi_{\phi\phi}$ , and heterotypic interactions,  $\chi_{\phi\psi}$ , may exhibit either segregative LLPS, driven by homotypic interactions, or associative LLPS, driven by heterotypic interactions. This motivates us to consider three classes of such ternary systems that relate to the transition between segregative to associative LLPS as  $\chi_{\phi\psi}$  is increased relative to  $\chi_{\phi\phi}$ . These three classes differ in the magnitude of their homotypic interactions: (1) Ternary systems in which the homotypic interaction  $\chi_{\phi\phi}$  is large enough so that the scaffold  $\phi$  can phase separate even in the absence of the multivalent binding partner  $\psi$ . This indicates that the binodals in the phase diagrams of this class intersect the  $\psi = 0$  line for any value of  $\chi_{\phi\psi}$ . Consequently, this class of systems can phase separate for any value of  $\chi_{\phi\psi}$ , so that the transition from segregative to associative LLPS is continuous as  $\chi_{\phi\psi}$  is increased. (2) Ternary systems in which the homotypic interaction  $\chi_{\phi\phi}$  is too small to promote phase separation in the absence of the multivalent binding partner  $\psi$ , so that their binodals do not intersect the line  $\psi = 0$ . However, for finite concentrations of  $\psi$ , the crowding due to the  $\psi$  may allow for phase separation. Below, we show that this class of systems may exhibit either segregative or associative LLPS (depending on the values of the heterotypic and homotypic interactions). The transition between the two as  $\chi_{\phi\psi}$  is varied involves reentrant behavior: for small values of  $\chi_{\phi\psi}$ , LLPS is observed. The region of LLPS then shrinks to zero as  $\chi_{\phi\psi}$  is increased (i.e., the system is in a single homogenous phase), but for large enough values, the phase diagram again contains a region of LLPS. This is termed reentrant behavior of the phase separation. (3) Ternary systems in which the homotypic interaction parameter is too small to promote segregative LLPS even in the presence of the  $\psi$ . However, these systems may still show associative LLPS when  $\chi_{\phi\psi}$  is large enough. Since there is no segregative LLPS in this case, there is no transition between segregative and associative LLPS. In the subsections below, we consider each of these three classes of ternary systems separately, and analytically calculate the critical values of  $\chi_{\phi\phi}$  that distinguish these classes, as well as the critical values of  $\chi_{\phi\psi}$  that are associated with the transition from segregative to associative LLPS.

### 5.3.1 Class (1) - Large homotypic interaction: $\phi$ can phase separate in the absence of $\psi$

For this class of systems, the homotypic interactions are large enough so that the scaffold  $\phi$  can phase separate in the absence of solute  $\psi$ . The free energy of the binary system that result from setting  $\psi = 0$  is the same as the one described in Eq. 47. Therefore, the critical homotypic interaction above which a ternary system belongs to class (1), is the same as  $\chi_{\phi\phi}^c$  defined in Eq. 50. In the limit  $\chi_{\phi\psi} \rightarrow 0$ , systems of case III are reduced to those of case I, and their phase diagrams in this limit must be similar to the one shown in Fig. 3A in the main text, namely ones in which the binodals intersect both the line  $\psi = 0$  and the line  $\phi + \psi = 1$ . However, the phase diagram shown in Fig. 5C in the main text suggests that, as  $\chi_{\phi\psi}$  increases, the area of the binodal shrinks and the binodal ceases to intersect the line  $\phi + \psi = 1$ . Therefore, there must be a critical heterotypic interaction parameter above which the binodals of class (1) do not intersect the line  $\phi + \psi = 1$ . Because the line  $\phi + \psi = 1$  represents a binary blend of solutes  $\psi$  and  $\phi$  without solvent, we denote this critical value as  $\chi_{\phi\psi}^b$ . To calculate the value  $\chi_{\phi\psi}^b$ , we write the energy of the blend,  $f_b(\phi)$ , as a function of  $\phi$  by substituting  $\psi = 1 - \phi$  and  $\chi_{\psi\psi} = 0$  into the free energy in Eq. 38, which results in:

$$f_b(\phi) v_0 = \frac{\phi}{N_\phi} \log(\phi) + \frac{(1-\phi)}{N_\psi} \log(1-\phi) - \frac{1}{2} \tilde{\chi} \phi^2 \quad (70)$$

where  $\tilde{\chi} = \chi_{\phi\phi} - 2\chi_{\phi\psi}$  is the effective (term proportional to  $\phi^2$ ) homotypic interaction energy of the  $\phi$  solute in the blend; terms linear in  $\phi$  were omitted since they contribute only constant terms to the total free energy that renormalize the chemical potential and do not affect the phase separation properties of the system. The binary blend can phase separate when the binodal of the ternary system intersects the line  $\phi + \psi = 1$ , and its effective homotypic interaction,  $\tilde{\chi}$ , is larger than its value at the critical point for LLPS of the binary blend. The solute concentration,  $\phi_c$ , and effective homotypic interaction,  $\tilde{\chi}_c$ , at the critical point of the blend are the solutions of the set of equations  $\partial^2 f_b / \partial \phi^2 = 0$  and  $\partial^3 f_b / \partial \phi^3 = 0$  [14], which are expressed as follows:

$$0 = \frac{1}{N_\phi \phi_c} + \frac{1}{N_\psi (1 - \phi_c)} - \tilde{\chi}_c \quad (71)$$

$$0 = -\frac{1}{N_\phi \phi_c^2} + \frac{1}{N_\psi (1 - \phi_c)^2} \quad (72)$$

The critical interaction  $\tilde{\chi}_c$  that solves these these equations is:

$$\tilde{\chi}_c = \left( \frac{1}{\sqrt{N_\phi}} + \frac{1}{\sqrt{N_\psi}} \right)^2 \quad (73)$$

Since the effective homotypic interaction energy of the blend is related to the homotypic and heterotypic interactions in the ternary system by  $\tilde{\chi} = \chi_{\phi\phi} - 2\chi_{\phi\psi}$ , it is larger than  $\tilde{\chi}_c$  if  $\chi_{\phi\phi} - \left( N_\phi^{-\frac{1}{2}} + N_\psi^{-\frac{1}{2}} \right)^2 > 2\chi_{\phi\psi}$ . This indicates that the critical value of the heterotypic interaction energy,  $\chi_{\phi\psi}^b$ , below which the binodal of a class (1) system intersects the line  $\phi + \psi = 1$ , is:

$$\chi_{\phi\psi}^b = \frac{1}{2} \chi_{\phi\phi} - \frac{1}{2} \left( \frac{1}{\sqrt{N_\phi}} + \frac{1}{\sqrt{N_\psi}} \right)^2 \quad (74)$$

When the heterotypic interaction  $\chi_{\phi\psi}$  is increased beyond  $\chi_{\phi\psi}^b$ , the LLPS changes continuously from segregative (negatively-sloped tie-lines) to associative (positively-sloped tie-lines). Therefore, a characteristic value of the heterotypic interaction, denoted by  $\chi_{\phi\psi}^t$ , for which the critical tie-line - i.e., the tangent to the critical point, is parallel to the  $\phi$  axis, must exist. In this particular case, the equations that determine the concentrations of solutes  $\phi$  and  $\psi$  at the critical point [26], are identical with those that determine the critical point in a binary system,  $\partial^2 f / \partial \phi^2 = 0$  and  $\partial^3 f / \partial \phi^3 = 0$  [14]. However, in this case the value of  $\chi_{\phi\phi}$  is fixed, and the unknowns critical parameters are  $\phi_c$  and  $\psi_c$ . Substitution of the free energy in Eq. 38 with



$\chi_{\psi\psi} = 0$  into these two simplified equations results in the following set of equations that characterizes the concentrations of the two solutes at the critical point,  $\phi_c$  and  $\psi_c$ :

$$\frac{1}{N_\phi\phi_c} + \frac{1}{(1-\phi_c-\psi_c)} - \chi_{\phi\phi} = 0 \quad (75)$$

$$-\frac{1}{N_\phi\phi_c^2} + \frac{1}{(1-\phi_c-\psi_c)^2} = 0 \quad (76)$$

whose solution is

$$\phi_c = \frac{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)}{\sqrt{N_\phi}\chi_{\phi\phi}} \quad (77)$$

$$\psi_c = 1 - \frac{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2}{\chi_{\phi\phi}} \quad (78)$$

These values of the concentrations at the critical point are correct only when the heterotypic interaction attains the characteristic value,  $\chi_{\phi\psi} = \chi_{\phi\psi}^t$ , which we now calculate. To relate the heterotypic interaction  $\chi_{\phi\psi}^t$  to the concentrations of the solutes at the critical point, we use the fact that the critical point must lie on the spinodal line [26]. Therefore, for  $\phi = \phi_c$ ,  $\psi = \psi_c$  (where  $\phi_c$  and  $\psi_c$  are respectively given by Eqs. 77 and 78), and  $\chi_{\phi\psi} = \chi_{\phi\psi}^t$ , the determinant of the Hessian  $H_{III}$  in this case, must vanish. For a ternary system of case III, where homotypic and heterotypic interactions are both present, the free energy is given by substitution of  $\chi_{\psi\psi} = 0$  in Eq. 38. The Hessian is then obtained by substitution of this free energy into Eq. 46, which results in the following expression:

$$H_{III} = \begin{pmatrix} \frac{1}{N_\phi\phi} + \frac{1}{1-\phi-\psi} - \chi_{\phi\phi} & \frac{1}{1-\phi-\psi} - \chi_{\phi\psi} \\ \frac{1}{1-\phi-\psi} - \chi_{\phi\psi} & \frac{1}{N_\psi\psi} + \frac{1}{1-\phi-\psi} \end{pmatrix} \quad (79)$$

When  $\chi_{\phi\psi} = \chi_{\phi\psi}^t$ , and the expressions from Eqs. 77 and 78 are used for  $\phi$  and  $\psi$ , setting  $\det(H_{III}) = 0$  results in an equation, whose solution gives the explicit expression of  $\chi_{\phi\psi}^t$  as a function of the solute sizes and the homotypic interaction:

$$\chi_{\phi\psi}^t = \frac{\chi_{\phi\phi}}{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)} \quad (80)$$

For  $\chi_{\phi\psi} < \chi_{\phi\psi}^t$ , the tie-lines are negatively-sloped (segregative LLPS), while for  $\chi_{\phi\psi} > \chi_{\phi\psi}^t$ , the tie-lines are positively-sloped (associative LLPS). Because an increase of the heterotypic interaction between  $\phi$  and  $\psi$  favor LLPS in which the concentrated phase has considerable amounts of both solutes, the heterotypic interactions destabilize segregative LLPS and stabilize associative LLPS, as indicated by their effect on the area of the binodals (see Fig. 5D in the main text). Therefore, the area of the binodal when the heterotypic interaction parameter is equal to the characteristic one,  $\chi_{\phi\psi} = \chi_{\phi\psi}^t$ , is expected to be minimal.

### 5.3.2 Class (2) - Intermediate homotypic interaction: $\phi$ can phase separate segregatively but requires crowding interactions

For ternary systems of this class, the homotypic interaction  $\chi_{\phi\phi}$  is too small to promote LLPS in binary systems of the solute  $\phi$  and a solvent, but is large enough to drive LLPS when assisted by the crowding interactions with finite concentrations of the multivalent binding partner  $\psi$ . Consequently, the binodals of systems in class (2) do not intersect the line  $\psi = 0$ . Therefore, in this case,  $\chi_{\phi\phi}$  is smaller than the critical homotypic interaction  $\chi_{\phi\phi}^c$  given by Eq. 50, but must be larger than a lower bound, below which,  $\phi$  cannot phase separate homotypically at all. Since, as explained at the end of the previous subsection, heterotypic

interactions of  $\phi$  with  $\psi$  inhibit segregative LLPS,  $\psi$  is most effective in promoting segregative LLPS when  $\chi_{\phi\psi} \rightarrow 0$ , namely when  $\psi$  is a crowder (with only excluded volume interactions with  $\phi$  and the solvent) rather than a multivalent binding partner (with non-zero attraction to  $\phi$  for  $\chi_{\phi\psi} > 0$ ). Therefore, the lower bound for the homotypic interactions below which the solute  $\phi$  cannot undergo segregative LLPS at all, even when promoted by crowding interactions of  $\psi$ , is the critical value  $\chi_{\phi\phi}^m$  derived for case I in Eq. 54;  $\chi_{\phi\phi}^m$  is the minimal (critical) homotypic interaction required for phase separation of the binary blend of  $\phi$  and  $\psi$  in the absence of solvent, i.e., along the line  $\phi + \psi = 1$ . In conclusion, the homotypic interactions of class (2) systems must lie within the range  $\chi_{\phi\phi}^c > \chi_{\phi\phi} > \chi_{\phi\phi}^m$ .

We note that this interaction range  $\chi_{\phi\phi}^c > \chi_{\phi\phi} > \chi_{\phi\phi}^m$  corresponds to type (2) phase diagrams of systems considered in case I (where  $\chi_{\phi\psi} = 0$ ), whose binodals intersect the line  $\phi + \psi = 1$  but not the line  $\psi = 0$  (Fig. 3B in the main text). Therefore, in the limiting case  $\chi_{\phi\psi} \rightarrow 0$ , the phase diagrams of class (2) systems must coincide with type (2) diagrams of case I. However, because heterotypic interactions destabilize segregative LLPS, an increase of  $\chi_{\phi\psi}$  from zero causes the binodal to shrink, and to eventually become a closed curve that does not intersect the line  $\phi + \psi = 1$  (see Fig. 5A in the main text). This is similar to the behavior of class (1) phase-diagrams, whose binodals no longer intersect the line  $\phi + \psi = 1$  when  $\chi_{\phi\psi}$  exceeds the critical value  $\chi_{\phi\psi}^b$ . Moreover, the derivation of  $\chi_{\phi\psi}^b$  for class (1) systems above, does not require the value of  $\chi_{\phi\phi}$  to be within the range that characterizes LLPS for class (1) systems. Therefore, the critical value  $\chi_{\phi\psi}^b$  given by Eq. 74, is also the critical heterotypic interaction below (above) which binodals of class (2) systems intersect (do not intersect) the line  $\phi + \psi = 1$ . Surprisingly, in contrast to class (1) phase-diagrams, a further increase of  $\chi_{\phi\psi}$  in class (2) phase diagrams causes the system to exhibit reentrant behavior. This is because the binodal loop of class (2) systems shrinks with increasing values of  $\chi_{\phi\psi}$  and eventually disappear when  $\chi_{\phi\psi}$  reaches a critical value  $\chi_{\phi\psi}^{lc}$  (see below for its derivation). For larger values of  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$  (see Fig. 5B in the main text and below for its derivation), the binodal loop reappears. The phase separation is now associative instead of segregative and the loop expands as  $\chi_{\phi\psi}$  is increased beyond the value of  $\chi_{\phi\psi}^{uc}$ . When the heterotypic interaction is in between these two critical values,  $\chi_{\phi\psi}^{lc} < \chi_{\phi\psi} < \chi_{\phi\psi}^{uc}$ , LLPS does not occur and the mixed phase of the ternary system is the equilibrium state for any concentrations of the solutes  $\phi$  and  $\psi$ .

To calculate these two critical values,  $\chi_{\phi\psi}^{lc}$  and  $\chi_{\phi\psi}^{uc}$ , as functions of the homotypic interaction  $\chi_{\phi\phi}$  and the solute sizes,  $N_\phi$  and  $N_\psi$ , we determine the values of the heterotypic interaction strength at which the spinodal vanishes and then reappears. As explained at the beginning of the ‘‘Supplementary results’’ section above, the binodal region must enclose a spinodal region, so that the values of  $\chi_{\phi\psi}$  for which the spinodal vanishes and reappears must be the same as  $\chi_{\phi\psi}^{lc}$  and  $\chi_{\phi\psi}^{uc}$ , respectively. Furthermore, since the binodal curves are closed loops prior to their vanishing and after their reappearance, the spinodal must also be a closed loop enclosed by the binodal.

The spinodal is defined by the condition  $\det(H_{III}) = 0$  (where  $H_{III}$  is given by Eq. 79), which is expressed in terms of the solutes concentrations, their sizes, and the two interaction parameters, as follows:

$$0 = N_\phi (\chi_{\phi\phi} + \chi_{\phi\psi}^2 N_\psi \psi) \phi^2 \quad (81)$$

$$+ ((N_\phi - 1) - N_\phi \chi_{\phi\phi} (1 + (N_\psi - 1) \psi) + N_\phi N_\psi \chi_{\phi\psi} \psi (2 - \chi_{\phi\psi} (1 - \psi))) \phi + (1 + (N_\psi - 1) \psi) \quad (82)$$

We now analyze the spinodal, in a manner similar to that used in subsection ‘‘Case II - Heterotypic ( $\chi_{\phi\phi} = 0, \chi_{\phi\psi} > 0$ ): Concentrated phase rich in both solutes (associative LLPS)’’. This spinodal is a quadratic function of  $\phi$ , whose discriminant,  $\Delta_{III}(\psi)$ , is:

$$\Delta_{III}(\psi) = ((N_\phi - 1) - N_\phi \chi_{\phi\phi} (1 + (N_\psi - 1) \psi) + N_\phi N_\psi \chi_{\phi\psi} \psi (2 - \chi_{\phi\psi} (1 - \psi)))^2 \quad (83)$$

$$- 4N_\phi (1 + (N_\psi - 1) \psi) (\chi_{\phi\phi} + \chi_{\phi\psi}^2 N_\psi \psi) \quad (84)$$

There are two real values of  $\phi$  that solve the quadratic equation 81, as long as the concentration  $\psi$  that appear in Eq. 81 gives a positive discriminant:  $\Delta_{III}(\psi) > 0$ . This means that for any value of  $\psi$  where  $\Delta_{III}(\psi) > 0$ , there are two points (due to the two solutions of the quadratic equation for the spinodal) on the spinodal that are associated with this value of  $\psi$ ; those two points generally have different values of  $\phi$ . Thus, continuous regions of  $\psi$  that satisfy the inequality  $\Delta_{III}(\psi) \geq 0$ , form closed curves in the  $\phi$ - $\psi$  plane,

whose maximal and minimal values of  $\psi$  are those that satisfy the equality  $\Delta_{III}(\psi) = 0$ . If these regions, and their corresponding values of  $\phi$ , (calculated from Eq. 81), are within the “physical” range of the two concentrations,  $\phi, \psi \geq 0$  and  $\phi + \psi \leq 1$ , they determine closed, spinodal curves. Below, we determine the roots of the equation  $\Delta_{III}(\psi) = 0$ , and calculate their corresponding values of  $\phi$  using Eq. 81. We then select the “physical” roots that correspond to the the maximal and minimal values of  $\psi$  along the spinodal, whose corresponding  $(\phi, \psi)$  points are in the region  $\phi, \psi \geq 0$  and  $\phi + \psi \leq 1$ . Since Eq. 83 is a quartic polynomial in  $\psi$ , it has four roots that can be analytically expressed as:

$$\psi_1 = \frac{1}{2} \left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right) \quad (85)$$

$$- \frac{1}{2} \sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)} \quad (86)$$

$$\psi_2 = \frac{1}{2} \left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right) \quad (87)$$

$$+ \frac{1}{2} \sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)} \quad (88)$$

$$\psi_3 = \frac{1}{2} \left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right) \quad (89)$$

$$- \frac{1}{2} \sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)} \quad (90)$$

$$\psi_4 = \frac{1}{2} \left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right) \quad (91)$$

$$+ \frac{1}{2} \sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)} \Rightarrow \quad (92)$$

where we used the Solve function of Wolfram Mathematica software to obtain these roots. The corresponding values of  $\phi$  that are calculated for these four values of  $\psi$  using Eq. 81 are:

$$\phi_1 = \frac{1}{2} \frac{1}{\sqrt{N_\phi} + 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \left( 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) \right) \quad (93)$$

$$+ \frac{1}{2} \frac{\sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)}}{\sqrt{N_\phi} + 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \quad (94)$$

$$\phi_2 = \frac{1}{2} \frac{1}{\sqrt{N_\phi} + 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \left( 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) \right) \quad (95)$$

$$- \frac{1}{2} \frac{\sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)}}{\sqrt{N_\phi} + 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \quad (96)$$

$$\phi_3 = -\frac{1}{2} \frac{1}{\sqrt{N_\phi} - 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \left( 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) \right) \quad (97)$$

$$- \frac{1}{2} \frac{\sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)}}{\sqrt{N_\phi} - 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \quad (98)$$

$$\phi_4 = -\frac{1}{2} \frac{1}{\sqrt{N_\phi} - 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \left( 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) \right) \quad (99)$$

$$+ \frac{1}{2} \frac{\sqrt{\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 - \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 - \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right)}}{\sqrt{N_\phi} - 1 - \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}} \sqrt{N_\phi}} \quad (100)$$

It is important to note that, because of the discontinuous, reentrant behavior of the phase-diagram, the two “physical” points  $(\phi, \psi)$  of the four pairs of solutions are not necessarily the same for the separate continuous domains  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$  and  $\chi_{\phi\psi} < \chi_{\phi\psi}^{lc}$ . We thus consider each domain separately and select the two points that are physical.

For values of  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$ , the binodal and spinodal form closed loops, whose maximal and minimal values of  $\psi$  and their corresponding values of  $\phi$ , are all in the physical region of the concentration space for any value of  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$  (see Fig. 5B in the main text). Therefore, we consider the asymptotic limit,  $\chi_{\phi\psi} \rightarrow \infty$  for which the expressions for the various sets of values of  $\psi$  and  $\phi$  detailed above are greatly simplified, and use these simplified expressions to select the values that are physical. In the limit  $\chi_{\phi\psi} \rightarrow \infty$ ,  $(\phi_1, \psi_1) \rightarrow \left( (\sqrt{N_\phi} + 1)^{-1}, 0^+ \right)$ ,  $(\phi_2, \psi_2) \rightarrow (0^+, 1^-)$ ,  $(\phi_3, \psi_3) \rightarrow \left( -(\sqrt{N_\phi} - 1)^{-1}, 0^+ \right)$  and  $(\phi_4, \psi_4) \rightarrow (0^-, 1^-)$ ;  $x^-$  and  $x^+$  respectively denote values that are infinitesimally smaller or larger than  $x$  (where the relevant  $x$  values are 0 or 1). This indicates that the points  $(\phi_1, \psi_1)$  and  $(\phi_2, \psi_2)$  are within the physical region of first quadrant of the plane  $(\phi, \psi)$ , while  $(\phi_3, \psi_3)$  and  $(\phi_4, \psi_4)$ , whose  $\phi$  values are negative, are unphysical.

For values of the heterotypic interaction  $\chi_{\phi\psi} < \chi_{\phi\psi}^{lc}$ , we consider the limiting case  $\chi_{\phi\psi} \rightarrow 0$  for which the expressions for the four points simplify, and use these simplified expressions to select the two physical points. However, we note that in contrast to the limit  $\chi_{\phi\psi} \rightarrow \infty$ , in the limit  $\chi_{\phi\psi} \rightarrow 0$ , (whose typical phase diagram is presented in Fig. 3B in the main text), the binodal does not form a closed loop; instead, its two branches intersect the line  $\phi + \psi = 1$ . Due to the relation of the binodal and spinodal, this suggests that the points along the spinodal corresponding to the minimal and maximal values  $\psi$ , are in the first quadrant  $\phi, \psi > 0$ , although not necessarily in the physical region of this quadrant,  $\phi + \psi \leq 1$ . In the limit  $\chi_{\phi\psi} \rightarrow 0$ , we find that the points  $(\phi_1, \psi_1)$  and  $(\phi_2, \psi_2)$  are in the first quadrant while  $(\phi_3, \psi_3)$  and  $(\phi_4, \psi_4)$  are not. Thus, the points  $(\phi_1, \psi_1)$  and  $(\phi_2, \psi_2)$  are the physical ones for the region  $\chi_{\phi\psi} < \chi_{\phi\psi}^{lc}$  as well as for  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$ .

When the heterotypic interaction strength is equal to either of the critical ones,  $\chi_{\phi\psi}^{lc}$  or  $\chi_{\phi\psi}^{uc}$ , the spinodal collapses to a single point where the two points of maximal and minimal  $\psi$  are identical. Therefore, equating the expressions for either the  $\phi$  value (Eqs. 93 and 95), or their corresponding  $\psi$  values (Eqs. 85 and 87) that give the two extremal points along the spinodal, provides an equation whose solutions are the critical heterotypic interactions  $\chi_{\phi\psi}^{lc}$  and  $\chi_{\phi\psi}^{uc}$ . The equation that results from equating either Eqs. 93 and 95, or Eqs. 85 and 87, is the same quartic equation of  $\chi_{\phi\psi}$ :

$$\left( 1 + \frac{\chi_{\phi\phi}}{\chi_{\phi\psi}^2} \left( 1 - \frac{1}{N_\psi} \right) - \frac{2}{\chi_{\phi\psi}} \left( 1 + \frac{1}{\sqrt{N_\phi}} \right) \right)^2 - \frac{4}{\chi_{\phi\psi}^2} \frac{1}{N_\psi} \left( \left( 1 + \frac{1}{\sqrt{N_\phi}} \right)^2 - \chi_{\phi\phi} \right) = 0$$

which has four roots. However, substituting two of these four roots into Eqs. 85 and 87 results in negative values of  $\psi$ , so that these two values  $\chi_{\phi\psi}$  cannot be the critical ones, since the spinodal point must lie in the physical region of concentration space,  $\phi, \psi \geq 0$  and  $\phi + \psi \leq 1$ . The two remaining values of  $\chi_{\phi\psi}$  are thus

the critical ones; as required by the definitions of  $\chi_{\phi\psi}^{lc}$  and  $\chi_{\phi\psi}^{uc}$ , the smaller one is  $\chi_{\phi\psi}^{lc}$  and the larger one is  $\chi_{\phi\psi}^{uc}$ , which are expressed as follows:

$$\chi_{\phi\psi}^{lc} = \left(1 - \frac{1}{\sqrt{N_\psi}}\right) \left( \left(1 + \frac{1}{\sqrt{N_\phi}}\right) - \sqrt{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2 - \chi_{\phi\phi}} \right) \quad (101)$$

$$\chi_{\phi\psi}^{uc} = \left(1 + \frac{1}{\sqrt{N_\psi}}\right) \left( \left(1 + \frac{1}{\sqrt{N_\phi}}\right) + \sqrt{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2 - \chi_{\phi\phi}} \right) \quad (102)$$

The two critical values for  $\chi_{\phi\psi}$  that we have found, determine the nature of LLPS in ternary systems of this class. For  $\chi_{\phi\psi} < \chi_{\phi\psi}^{lc}$ , the LLPS is segregative, while for  $\chi_{\phi\psi} > \chi_{\phi\psi}^{uc}$ , the LLPS is associative; in between, when  $\chi_{\phi\psi}^{uc} > \chi_{\phi\psi} > \chi_{\phi\psi}^{lc}$ , the system does not exhibit LLPS and the mixed phase is the equilibrium state for all values of the concentrations  $\phi$  and  $\psi$ .

### 5.3.3 Class (3) - Small homotypic interaction: $\phi$ can only phase separate associatively for finite values of $\psi$

For ternary systems of this class, the homotypic interactions  $\chi_{\phi\phi}$  are too small to facilitate segregative LLPS even when supplemented by crowding interactions. This indicates that, even if the multivalent binding partner is replaced with a crowder (i.e., one that has only excluded volume interactions with the other solute and the solvent, but no heterotypic interactions), the system will not phase separate. As explained in the subsection ‘‘Case I - Homotypic ( $\chi_{\phi\phi} > 0, \chi_{\phi\psi} = 0$ ): Concentrated phase rich in one solute (segregative LLPS)’’, this means that the homotypic interaction strength is less than  $\chi_{\phi\phi}^m$ , which is specified in Eq. 54. Since heterotypic interactions inhibit segregative LLPS, systems of this class, which do not exhibit segregative LLPS for  $\chi_{\phi\psi} = 0$  (i.e., where  $\psi$  is a crowder), will not do so also in the presence of a positive heterotypic interaction  $\chi_{\phi\psi} > 0$ . However, such systems can still undergo associative LLPS if the heterotypic interaction  $\chi_{\phi\psi}$  is large enough. Below, we determine the critical heterotypic interaction above which, class (3) systems undergo associative LLPS.

The critical value of the heterotypic interaction can be calculated from the spinodal curve. Since the spinodal and binodal must both exist when one of them exists, the critical heterotypic interaction for which the spinodal appears is the one above which the system phase separates associatively. Moreover, since the Hessians of classes (2) and (3) are identical, the critical interaction parameters  $\chi_{\phi\psi}^{lc}$  and  $\chi_{\phi\psi}^{uc}$  respectively given by Eqs. 101 and 102, mark the disappearance and appearance of the spinodal for this class as well. The lower critical parameter,  $\chi_{\phi\psi}^{lc}$ , does not give rise to a physical transition for class (3) systems. This is because the concentrations of solutes,  $\phi$  and  $\psi$ , at the spinodal point for  $\chi_{\phi\psi} = \chi_{\phi\psi}^{lc}$ , (obtained by substitution of  $\chi_{\phi\psi}^{lc}$  from Eq. 101 into Eqs. 85 and 93), are non-physical concentrations that satisfy  $\phi + \psi > 1$  when  $\chi_{\phi\phi} < \chi_{\phi\phi}^m$ . In contrast, the upper critical interaction  $\chi_{\phi\psi}^{uc}$ , which marks the onset of associative LLPS in class (2) systems is physically relevant for class (3) systems and is the critical parameter above which the systems undergo associative LLPS. It is important to note that in the limit of vanishing homotypic interactions  $\chi_{\phi\phi} \rightarrow 0$ , the critical heterotypic interaction  $\chi_{\phi\psi}^{uc}$ , given by Eq. 102, coincides with the critical interaction  $\chi_{\phi\psi}^c$ , given by Eq. 69 for systems of case II. This is consistent with the results above since for  $\chi_{\phi\phi} \rightarrow 0$ , systems of case III are equivalent to those of case II. Furthermore, the definitions of  $\chi_{\phi\psi}^c$  and  $\chi_{\phi\psi}^{uc}$  imply that  $\chi_{\phi\psi}^c > \chi_{\phi\psi}^{uc}$  for any  $\chi_{\phi\phi} > 0$ , which suggests that even weak homotypic interactions stabilize associative LLPS, driven by heterotypic interactions. This is consistent with the results (Fig. 5D) and explanations (see subsection ‘‘Heterotypic interactions: in the absence or together with homotypic ones’’) provided in the main text.

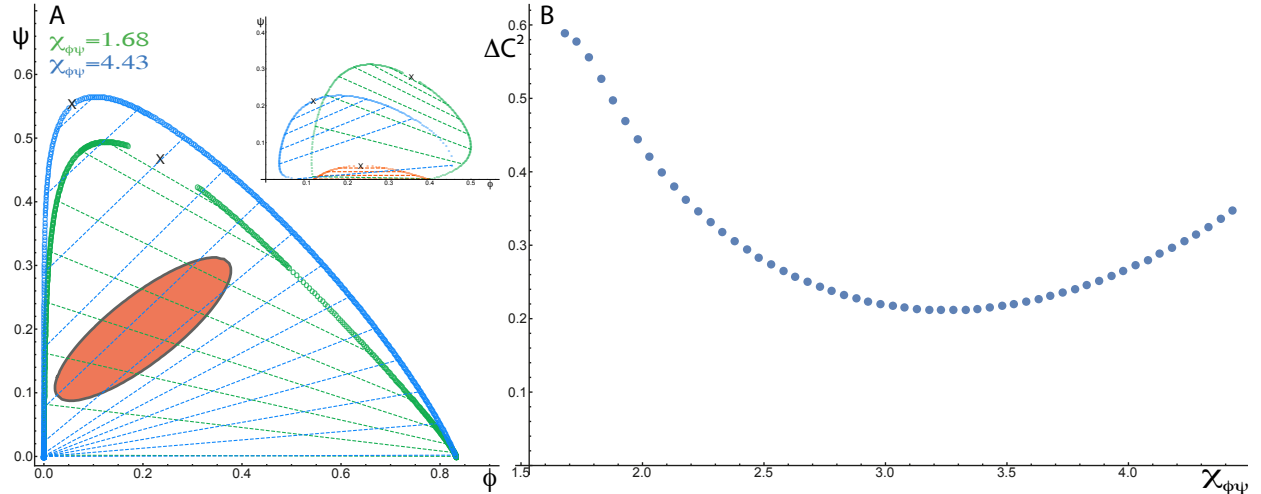


Figure 1: Effect of heterotypic interaction strength on magnitude of the noise in the solute concentrations at the coexisting domains. (A) Phase diagrams, tie-lines, and critical points (shown by black “X” marks) of two ternary systems consisting of multivalent solute  $\phi$  and multivalent binding partner  $\psi$  in a solvent, similar to those shown in Fig. 5C in the main text (shown in inset); the same noise ellipse characterizes the distribution of overall solute concentrations for both systems and is drawn as an orange ellipse. The lengths of solutes and homotypic interactions in the two systems are the same:  $N_\phi = 10$ ,  $N_\psi = 8$ , and  $\chi_{\phi\phi} = 3$ . However, in one system the heterotypic interactions are weak ( $\chi_{\phi\psi} \approx 1.68$ ) and the phase separation is segregative (green diagram), while in the other system the heterotypic interactions are strong ( $\chi_{\phi\psi} \approx 4.43$ ) and the phase separation is associative (blue diagram). The axes of the noise ellipse correspond to noisy expression of a pair of genes with intrinsic noise of 0.19 and extrinsic noise of 0.32 (similar to wild type *lacI*<sup>+</sup> in MG22 *E. coli* [19]), and its tilt angle is 30°. The slope of the noise ellipse is larger than the negatively sloped tie-lines of the green diagram but smaller than the positively sloped tie-lines of the blue diagram. The center of the noise ellipse was chosen to be at unrealistically large volume fractions ( $\phi = 0.2$ ,  $\psi = 0.2$ ) to allow easy visualization. (B) The total noise in the dilute phase as function of the heterotypic interaction for values of the heterotypic interaction ranging between those of the two systems presented in (A), spaced by steps of 0.05  $k_B T$ . Similar to [27], the total noise  $\Delta C^2$  is defined as the sum, for the two solutes, of the mean square deviations of the solute concentrations from their average concentration in the absence of noise, divided by the average concentration squared,  $\Delta C^2 = \langle (\phi_d - \langle \phi_d \rangle)^2 \rangle / \langle \phi_d \rangle^2 + \langle (\psi_d - \langle \psi_d \rangle)^2 \rangle / \langle \psi_d \rangle^2$ , where  $\phi_d$  and  $\psi_d$  are the solute concentrations in the dilute phase.  $\Delta C^2$  includes additional contributions that do not depend only on the alignment of the orientations of the noise ellipse and the tie-lines, such as the splay of the tie-lines or the increase/decrease of the average solute concentrations. Nonetheless, the plot of the noise,  $\Delta C^2$ , against the heterotypic interaction strength,  $\chi_{\phi\psi}$ , shows that the total noise decreases with increasing heterotypic interaction as the tie-lines become more aligned with the noise ellipse, and eventually increases as the tie-lines become misaligned again with the noise ellipse. Although the system is phase-separated for all the plotted values of heterotypic interactions, the quality of concentration buffering greatly varies, with the noise almost three-fold larger when the tie-lines are completely misaligned (green diagram in A) compared to the minimum of the noise in B.

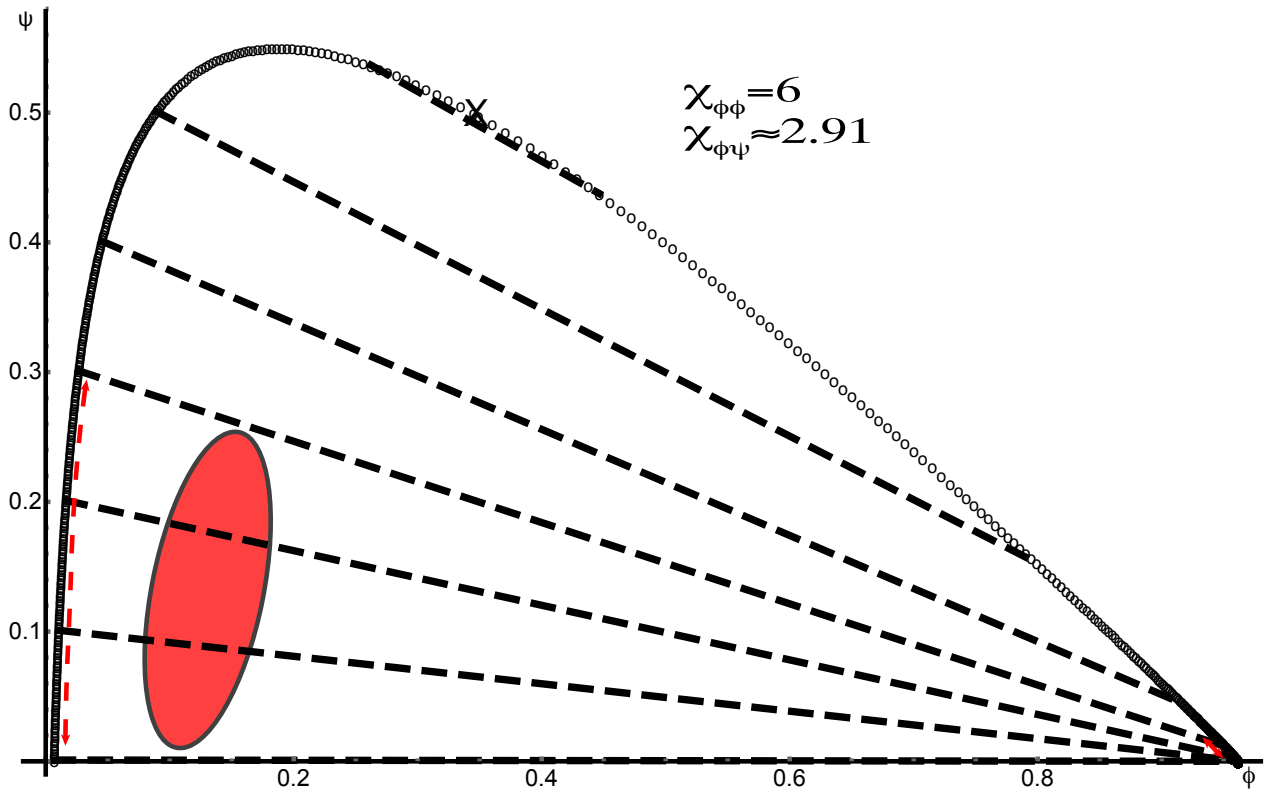


Figure 2: A ternary system that exhibits a phase diagram with highly splayed tie-lines. Binodal points are plotted as black, open circles, tie-lines as black, dashed lines, and the critical point with a black “X” mark. To highlight the possible concentration buffering capacities of such splayed phase diagram, an example of a noise-ellipse that is almost perpendicular to the tie-lines is plotted in red. The projected variations of the concentrations in the two coexisting phases due to the noise in the overall concentrations are indicated by red, dashed double headed arrows. Due to the large mismatch between the orientations of the noise-ellipse and the tie-lines, there is a wide range of possible equilibrium concentrations, so that conventional concentration buffering that buffers noise in both dilute and dense phases is not effective. However, as can be seen, the large splay of the tie-lines renders the dense phase robust to the large noise in the overall concentrations, while it augments the effect of the noise on the dilute phase. The parameters used to generate this phase diagram are  $N_\phi = 2$ ,  $N_\psi = 15$ ,  $\chi_{\phi\phi} = 6$ , and  $\chi_{\phi\psi} \approx 2.91$ .

Critical interaction symbol	Full expression as a function of the relevant lengths and interactions
$\chi_{\phi\phi}^c$	$\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2$
$\chi_{\phi\phi}^m$	$\left(\frac{1}{\sqrt{N_\phi}} + \frac{1}{\sqrt{N_\psi}}\right)^2$
$\chi_{\phi\psi}^c$	$2\left(1 + \frac{1}{\sqrt{N_\psi}}\right)\left(1 + \frac{1}{\sqrt{N_\phi}}\right)$
$\chi_{\phi\psi}^b$	$\frac{1}{2}\chi_{\phi\phi} - \frac{1}{2}\left(\frac{1}{\sqrt{N_\phi}} + \frac{1}{\sqrt{N_\psi}}\right)^2$
$\chi_{\phi\psi}^t$	$\chi_{\phi\phi} / \left(1 + \frac{1}{\sqrt{N_\phi}}\right)$
$\chi_{\phi\psi}^{lc}$	$\left(1 - \frac{1}{\sqrt{N_\psi}}\right)\left(\left(1 + \frac{1}{\sqrt{N_\phi}}\right) - \sqrt{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2 - \chi_{\phi\phi}}\right)$
$\chi_{\phi\psi}^{uc}$	$\left(1 + \frac{1}{\sqrt{N_\psi}}\right)\left(\left(1 + \frac{1}{\sqrt{N_\phi}}\right) + \sqrt{\left(1 + \frac{1}{\sqrt{N_\phi}}\right)^2 - \chi_{\phi\phi}}\right)$

Table 1: Explicit, analytical expressions of the critical interactions energies derived in ‘‘Supplementary results’’

Topological property	lower bound on $\chi_{\phi\phi}$ value	upper bound on $\chi_{\phi\phi}$ value	lower bound on $\chi_{\phi\psi}$ value	upper bound on $\chi_{\phi\psi}$ value
binodal intersects the line $\psi = 0$	$\chi_{\phi\phi}^c$	$\infty$	Independent	Independent
binodal intersects the line $\phi + \psi = 1$	$\chi_{\phi\phi}^m$	$\infty$	0	$\chi_{\phi\psi}^b$
Segregative LLPS: Continuous transition	$\chi_{\phi\phi}^c$	$\infty$	0	$\chi_{\phi\psi}^t$
Associative LLPS: Continuous transition	$\chi_{\phi\phi}^c$	$\infty$	$\chi_{\phi\psi}^t$	$\infty$
Segregative LLPS: Reentrant transition	$\chi_{\phi\phi}^m$	$\chi_{\phi\phi}^c$	0	$\chi_{\phi\psi}^{lc}$
Associative LLPS: Reentrant transition	0	$\chi_{\phi\phi}^c$	$\chi_{\phi\psi}^{uc}$	$\infty$

Table 2: Classification of phase-diagram topologies: For each topological property considered in the section ‘‘Supplementary results’’, we specify the ranges of the homotypic and heterotypic interactions (lower and upper bounds) for which the given property applies. Motivated by the absence of specific repulsive interactions (above and beyond excluded volume, which for a screening length of 1nm also subsumes the Coulomb repulsion of like-charged moieties) in biological systems, we limit all interactions to be attractive (positive parameters). The terms ‘‘Continuous’’ and ‘‘Reentrant’’ describe the two possible types of transitions between segregative and associative LLPS, as we have analyzed here. Intersection of the binodal with the line  $\psi = 0$  is independent of the heterotypic interaction because in the absence of solute  $\psi$ , the heterotypic interactions have no effect.



## References

- [1] James S Langer. Theory of spinodal decomposition in alloys. *Annals of Physics*, 65(1):53–86, 1971.
- [2] Sharon C Glotzer, Edmund A Di Marzio, and M Muthukumar. Reaction-controlled morphology of phase-separating mixtures. *Physical review letters*, 74(11):2034, 1995.
- [3] Ron Milo and Rob Phillips. *Cell biology by the numbers*. Garland Science, 2015.
- [4] Roy D Dar, Brandon S Razooky, Abhyudai Singh, Thomas V Trimeloni, James M McCollum, Chris D Cox, Michael L Simpson, and Leor S Weinberger. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*, 109(43):17454–17459, 2012.
- [5] Christopher KRT Jones. Geometric singular perturbation theory. In *Dynamical systems*, pages 44–118. Springer, 1995.
- [6] Yuting I Li and Michael E Cates. Non-equilibrium phase separation with reactions: a canonical model and its behaviour. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(5):053206, 2020.
- [7] Jacob J Christensen, Ken Elder, and Hans C Fogedby. Phase segregation dynamics of a chemically reactive binary mixture. *Physical Review E*, 54(3):R2212, 1996.
- [8] Christoph A Weber, David Zwicker, Frank Jülicher, and Chiu Fan Lee. Physics of active emulsions. *Reports on Progress in Physics*, 82(6):064601, 2019.
- [9] Joshua A Riback, Daniel SW Lee, Jorine M Eeftens, David W Sanders, Lien Beckers, and Clifford P Brangwynne. Spatially constrained transcriptional sources drive asphericity of intracellular liquids. *Biophysical Journal*, 120(3):278a, 2021.
- [10] Amy Novick-Cohen and Lee A Segel. Nonlinear aspects of the cahn-hilliard equation. *Physica D: Nonlinear Phenomena*, 10(3):277–298, 1984.
- [11] Omar Adame-Arana and Samuel A Safran. Confined polymers in a poor solvent: The role of bonding to the surface. *Macromolecules*, 2021.
- [12] David Zwicker, Anthony A Hyman, and Frank Juelicher. Suppression of ostwald ripening in active emulsions. *Physical Review E*, 92(1):012317, 2015.
- [13] Clifford P Brangwynne, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935):1729–1732, 2009.
- [14] Samuel Safran. *Statistical thermodynamics of surfaces, interfaces, and membranes*. CRC Press, 2018.
- [15] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002.
- [16] Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.
- [17] Eric W. Weisstein. "first-order ordinary differential equation". mathworld- a wolfram web resource.
- [18] Eric W. Weisstein. "ellipse". from mathworld- a wolfram web resource. <https://mathworld.wolfram.com/ellipse.html>.
- [19] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

- [20] Sheng Mao, Derek Kuldinow, Mikko P Haataja, and Andrej Košmrlj. Phase behavior and morphology of multicomponent liquid mixtures. *Soft matter*, 15(6):1297–1311, 2019.
- [21] Marina Feric, Nilesh Vaidya, Tyler S Harmon, Diana M Mitrea, Lian Zhu, Tiffany M Richardson, Richard W Kriwacki, Rohit V Pappu, and Clifford P Brangwynne. Coexisting liquid phases underlie nucleolar subcompartments. *Cell*, 165(7):1686–1697, 2016.
- [22] Miroslav Fiedler. *Matrices and graphs in geometry*. Number 139. Cambridge University Press, 2011.
- [23] Dan Deviri and Sam A Safran. Equilibrium size distribution and phase separation of multivalent, molecular assemblies in dilute solution. *Soft Matter*, 2020.
- [24] Pierre-Gilles De Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.
- [25] Kurt Binder and Peter Fratzl. Spinodal decomposition. *Phase transformations in materials*, pages 409–480, 2001.
- [26] Saroj Kumar Nandi, Meta Heidenreich, Emmanuel D Levy, and Samuel A Safran. Interacting multivalent molecules: affinity and valence impact the extent and symmetry of phase separation. *arXiv preprint arXiv:1910.11193*, 2019.
- [27] A Klosin, F Oltsch, T Harmon, A Honigmann, F Jülicher, AA Hyman, and C Zechner. Phase separation provides a mechanism to reduce noise in cells. *Science*, 367(6476):464–468, 2020.