Phase separation provides a mechanism to drive phenotype switching

Lijun Hong,* Zhenquan Zhang,* Zihao Wang, Xiaochen Yu, and Jiajun Zhang†
Guangdong Province Key Laboratory of Computational Science, Sun Yat-sen University, Guangzhou 510275, People's Republic of China and School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China

(Received 18 March 2024; accepted 5 June 2024; published 28 June 2024)

Phenotypic switching plays a crucial role in cell fate determination across various organisms. Recent experimental findings highlight the significance of protein compartmentalization via liquid-liquid phase separation in influencing such decisions. However, the precise mechanism through which phase separation regulates phenotypic switching remains elusive. To investigate this, we established a mathematical model that couples a phase separation process and a gene expression process with feedback. We used the chemical master equation theory and mean-field approximation to study the effects of phase separation on the gene expression products. We found that phase separation can cause bistability and bimodality. Furthermore, phase separation can control the bistable properties of the system, such as bifurcation points and bistable ranges. On the other hand, in stochastic dynamics, the droplet phase exhibits double peaks within a more extensive phase separation threshold range than the dilute phase, indicating the pivotal role of the droplet phase in cell fate decisions. These findings propose an alternative mechanism that influences cell fate decisions through the phase separation process. As phase separation is increasingly discovered in gene regulatory networks, related modeling research can help build biomolecular systems with desired properties and offer insights into explaining cell fate decisions.

DOI: 10.1103/PhysRevE.109.064414

I. INTRODUCTION

Phenotypic switching, intricately linked with the bistability and bimodality of gene expression levels, plays a pivotal role in determining cell fate [1–3]. Bistability occurs in a wide range of biological systems, and deterministic models are frequently employed to analyze system properties in terms of regulatory mechanisms and dynamic parameters [4–8]. However, random fluctuations in gene expression are inevitable due to the randomness of chemical reactions and the low number of genes and ribonucleic acids (RNAs) in cells [9]. Then the patterns associated with bistability appear as a bimodal probability distribution of gene expression levels (the distribution of gene products that have two maxima). Bimodal distributions provide evidence for phenotypic switching [10], exemplified by phenomena such as the human immunodeficiency virus (HIV) latent-vs-active decision and the development of genetic competence in *Bacillus subtilis* [11,12], which are crucial for enhancing cell survival in fluctuating environments [13–15]. Therefore, it is essential to elucidate the mechanism underlying phenotypic transformation.

In recent years, the bistability and bimodality of gene expression products have primarily been associated with the dynamic control of transcriptional regulatory networks. The most common mode of regulation is autoregulation [16–22]. Autoregulation is a direct regulation of gene expression by corresponding gene products, controlling random fluctuations in protein concentration according to the state of the organism, thereby driving cell fate in a particular direction [17,23–25]. Positive feedback loops frequently give rise to bistability and bimodality by promoting further expression of the same gene [4,26–31]. Transcription factors (TFs) with positive feedback are crucial for fate determination, such as TBX1 during heart development [32]; HNF1A and HNF4A during pancreas development [33]; GATA2, RUNX1, and TCF4 during hematopoiesis [34,35]; and POU3F2, RORB, SATB2, and TBR1 during brain development [36,37].

On the other hand, cellular compartments can serve as mechanisms to regulate gene expression, leading to multiple stabilization of expression products [38–40]. Most compartments in eukaryotic cells are membrane bound, although there are exceptions. These membraneless compartments are formed by liquid-liquid phase separation (LLPS). Increasing evidence suggests that LLPS plays a crucial role in regulating gene expression. LLPS is a physical mechanism that can effectively regulate cellular fluctuations by storing excess gene expression products in the dense phase and releasing them into the dilute phase when product levels decrease [41,42]. This has led to LLPS being recognized as a faster mechanism of cellular regulation than transcriptional and translational control, and the resulting biomolecular condensates provide unique methods to control the biochemical environment of cells [43,44]. In previous research, we have found that increasing burst frequency can weaken the ability of phase separation to reduce noise, while burst size has a nonmonotonic relationship with this ability [45]. However, the mechanism of whether and how phase separation mediates complex phenotypic transformation remains unclear. This phenomenon of biological phase separation has become a paradigm for

*These authors contributed equally to this paper.
†Contact author: zhjiajun@mail.sysu.edu.cn

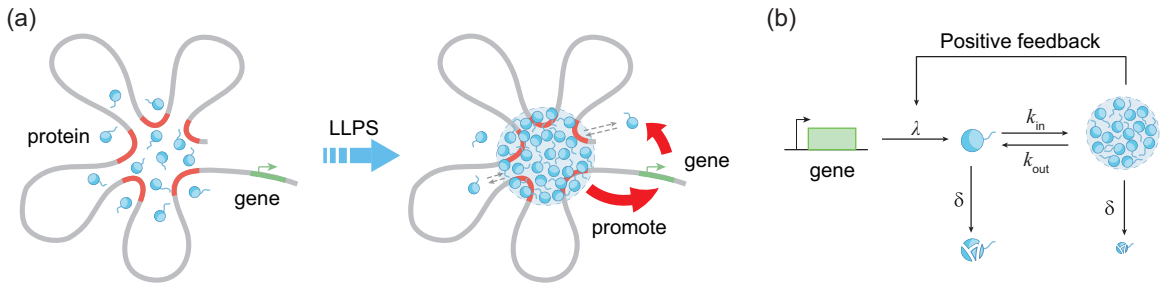

FIG. 1. A framework for analysis of transcriptional regulation model with phase separation. (a) Proteins are separated by liquid-liquid phase separation (LLPS) into a dilute phase and a droplet phase. The proteins in the droplet phase promote gene expression through positive feedback. (b) Gene reaction network of conjugation system with the phase separation process and the autoregulatory positive feedback process. The LLPS process is described dynamically by the random exchange of proteins between the dilute phase and the droplet phase.

understanding the intracellular organization of eukaryotes and bacteria [42,46–50].

Biologists have discovered that phase-separated assemblies of RNA polymerase II (RNA Pol II) and TFs could provide a general biophysical basis for gene expression processes. RNA-binding proteins, including TDP-43, FUS, and hnRNP A1, form phase-separated condensate and bind to the C-terminal domain (CTD) of RNA Pol II to participate in transcriptional initiation [51]. TFs have a deoxyribonucleic acid (DNA)-binding domain (DBD) with a specific structure and a transactivation domain (TAD) with an amorphous structure, where the TAD can interact with Mediator to form phase-separated condensates that activate gene expression [52,53]. The transcriptional coactivator Yes-associated protein (YAP) condensates enriched the transcription-related factors and subsequently induced the transcription process of YAP-specific proliferation genes [54,55]. Furthermore, phase separation can mediate the super-enhancer (SE) regulation of transcription [56]. Experiments have demonstrated that transcriptional coactivators BRD4 and MED1 form a condensate at the SEs [57]. These pieces of evidence indicate that LLPS is prevalent in the gene expression process.

Autoregulation and LLPS are important regulatory mechanisms in gene expression. However, the role of phase separation in autoregulatory networks is unclear and is critical to understanding cell fate decisions. In this paper, we explore the role of the synergistic behavior of phase separation and positive feedback during gene expression on biological fate decisions. Firstly, we introduce the LLPS process to autoregulated gene expression using a stochastic switching model to capture protein movement between dilute and droplet phases [41]. Then we establish reaction rate equations and chemical master equations (CMEs) for the system to explicitly capture the effect of LLPS on the behavior of the system. Bistability and bimodal properties of gene expression products can be induced in the dilute phase and the droplet phase, and bistable regions exist for the phase separation parameter. The double peaks of the dilute phase and the droplet phase do not appear simultaneously due to the stochastic dynamics. This paper is organized as follows. Section II provides a detailed description of the coupled model and obtains the rate equation in the deterministic case and the master equation in the stochastic case. Section III examines the bifurcation behavior of the deterministic system and the switching behavior of the stochastic system in the steady state. Section IV concludes this paper by discussing our results and their applicability.

II. MATERIALS AND METHODS

A. Biochemical reaction system with phase separation and stochastic gene expression

We proposed a model to investigate the role of LLPS in autoregulation. The model comprises two processes: the LLPS process of proteins and the stochastic gene expression process with feedback (Fig. 1). The LLPS process results in the formation of a dilute phase and a droplet phase, without affecting the total number of proteins; however, proteins can shuttle between these two distinct phases. The gene expression process autonomously regulates protein production, thereby altering the number of proteins in the dilute phase.

First, we construct a phase separation model in a binary mixture containing a solvent and a type of protein as solute to characterize the LLPS process. This simplified binary setup does not lose the essential property of reaction kinetics. We consider an incompressible two-phase system with a total volume of $V_{\rm tot}$ and a total protein number of $n_{\rm tot}$. One is the dilute phase (protein$_+$) with volume V_+ and protein number n_+, and the other is the condensed phase (protein$_-$) with volume V_- and protein number n_-. The incompressible system needs to satisfy conditions $V_+ + V_- = V_{\rm tot}$ and $n_+ + n_- = n_{\rm tot}$. Because the condensed phase exhibits liquidlike properties, it is also called a *droplet*. The droplet has a highly dynamic internal structure that allows rapid and stochastic material exchange with the dilute phase on short-time scales. The exchange rates $k_{\rm in}(n_+, n_-)$ and $k_{\rm out}(n_+, n_-)$ depend on the number of proteins in the two phases.

Second, we use a birth-and-death model to describe the gene expression process. The proteins generated by the birth-and-death process are in the intranuclear dilute phase. This is because proteins produced by translation return to the nucleus in a dissociated form. The protein production rate is not a constant [58] but a function of the number of protein molecules in the droplet phase. With the accumulation of n_+, $k_{\rm in}(n_+, n_-)$ and $k_{\rm out}(n_+, n_-)$ will be altered [Fig. 1(a)]. We assume that the protein can be degraded in both phases at a rate δ, and $\tau_p = 1/\delta$ can be used to characterize the time scale of gene expression.

Therefore, we introduce a stochastic model of the birth-and-death process with phase separation and feedback, which consists of the following five reactions [referring to Fig. 1(b)]:

$$\begin{aligned}
&\mathrm{DNA} \xrightarrow{\tilde{\lambda}(n_-)} \mathrm{DNA} + \mathrm{protein}_+,\\
&\mathrm{protein}_+ \underset{k_{\mathrm{out}}(n_+,n_-)}{\overset{k_{\mathrm{in}}(n_+,n_-)}{\rightleftharpoons}} \mathrm{protein}_-,\\
&\mathrm{protein}_+ \xrightarrow{\delta n_+} \emptyset,\\
&\mathrm{protein}_- \xrightarrow{\delta n_-} \emptyset,
\end{aligned} \tag{1}$$

where the protein synthesis occurs exclusively in the dilute phase, primarily because this process takes place on ribosomes located in the cytoplasm of eukaryotic cells. Synthesis depends on ribosomes having access to mRNA and amino acids to facilitate translation [59]. However, the aggregation of macromolecules within the droplet phase can restrict this accessibility, potentially impairing the efficiency of protein synthesis, while degradation of protein occurs in the dilute and droplet phases. We assume that protein production is subject to positive feedback from proteins in the droplet phase, the empirical dependence of which is well captured by a monotonically increasing form of a Hill-type function [60,61]:

$$\tilde{\lambda}(n_-) = \lambda + \frac{\zeta n_-^h}{K^h + n_-^h}, \tag{2}$$

where $\lambda > 0$ is the minimal or basal rate of protein generation, $\zeta > 0$ represents the feedback strength, the model degenerates into a process without feedback when $\zeta = 0$, $h > 0$ is the Hill coefficient, and $K > 0$ is the dissociation coefficient.

Experimental studies indicate that transcriptional bursts, which occur on a time scale of hours [62], contrast with phase separation processes, typically occurring on shorter time scales ranging from seconds to minutes [50,63–65]. The dynamics of phase separation exhibit significantly faster velocities than those of gene expression processes, so the model is divided into two parts using a time scale separation method. In Sec. II B, phase separation quickly reaches equilibrium on short time scales, so the rapid dynamics of phase separation are described by conditional probability distributions at fixed protein levels [66,67]. On a longer time scale, protein levels change with gene expression. Then phase separation quickly reaches a new equilibrium. In Sec. II C, we consider the effect of LLPS on protein distribution with feedback.

B. Stochastic dynamics of phase-separating system with fixed protein numbers

In this section, we consider a sufficiently short time scale during which there is no change in the amount of protein. At this point, the system exhibits only rapid Brownian motion, which allows it to quickly reach thermodynamic equilibrium. To describe an LLPS model with a fixed number of proteins, we define the effective volume fraction (concentration) of protein as $\phi_\pm = v n_\pm / V_\pm$, where v is the volume of a single molecule. Here, we consider the case of the droplet phase that is completely filled with protein molecules, so that the concentration of protein in the droplet phase is $\phi_- \equiv 1$. When the volume fraction of the system exceeds a certain threshold ϕ^*, it is expected to exhibit two distinct phases at a macroscopic level. In this section, we start from the macroscopic phase-separating systems at equilibrium, derive the explicit expression of ϕ^*, then give the mesoscopic details of the protein exchange between the two phases in reaction Eq. (1).

The equilibrium thermodynamics of the binary mixture system is described by minimizing the total free energy subject to the conservation of the number of protein molecules and their molecular volumes. We assume that the proteins in the dilute phase and droplet phase are homogeneous; then the interface energy is negligible. For simplicity, we ignore the interaction between other molecules, meaning that the free energy density is only related to entropy. Thus, the free energy densities of the dilute and droplet phases are $f(\phi_+) = k_\mathrm{B}T(\phi_+ \ln \phi_+ - \phi_+)/v$ and $f(\phi_-) = -\vartheta/v$, respectively [41], where T is the temperature, k_B denotes the Boltzmann constant, and ϑ is the relative chemical potential. Hence, the total free energy can be written as

$$\begin{aligned}
F(\phi_+, \phi_-) &= V_- f(\phi_-) + V_+ f(\phi_+)\\
&= -V_- \frac{\vartheta}{v} + V_+ \frac{k_\mathrm{B}T}{v}(\phi_+ \ln \phi_+ - \phi_+).
\end{aligned} \tag{3}$$

Equivalently, the free energy can be expressed in terms of the number of proteins as $F(n_+, n_-) = -\vartheta n_- + k_\mathrm{B}T n_+ \{\ln[n_+/(V_\mathrm{tot}/v - n_-)] - 1\}$.

By minimizing the total free energy of two coexisting phases with respect to constraints of constant volume and protein number, the threshold volume fraction for phase separation will be obtained. Given the total volume $V_+ + V_- = V_\mathrm{tot}$ and mass conservation condition $\phi_+ V_+ + \phi_- V_- = \phi_\mathrm{tot} V_\mathrm{tot}$, the equilibrium state requires minimizing the total free energy $F(\phi_+, \phi_-)$ [68]. Therefore, the following Lagrange function is constructed to satisfy these conditions:

$$\begin{aligned}
L = &-V_- \frac{\vartheta}{v} + V_+ \frac{k_\mathrm{B}T}{v}(\phi_+ \ln \phi_+ - \phi_+)\\
&- \mu[V_+\phi_+ + V_- - V_\mathrm{tot}\phi_\mathrm{tot}] + \pi[V_+ + V_- - V_\mathrm{tot}],
\end{aligned}$$

where μ and π are Lagrange multipliers. This results in the following equations:

$$\begin{aligned}
\frac{\partial L}{\partial V_+} &= \frac{k_\mathrm{B}T}{v}(\phi_+ \ln \phi_+ - \phi_+) - \mu\phi_+ + \pi = 0,\\
\frac{\partial L}{\partial V_-} &= -\frac{\vartheta}{v} - \mu + \pi = 0,\\
\frac{\partial L}{\partial \phi_+} &= V_+ \frac{k_\mathrm{B}T}{v} \ln \phi_+ - \mu V_+ = 0,\\
\frac{\partial L}{\partial \mu} &= V_+\phi_+ + V_- - V_\mathrm{tot}\phi_\mathrm{tot} = 0,\\
\frac{\partial L}{\partial \pi} &= V_+ + V_- - V_\mathrm{tot} = 0.
\end{aligned}$$

From the first three equations, we can obtain the critical threshold of protein concentration for the coexistence of the

two phases as

$$\phi^* = -W\left[-\exp\left(-\frac{\vartheta}{k_B T}\right)\right], \tag{4}$$

where W is the Lambert W function, the inverse function for any complex number [69], and ϕ^* is the concentration of the dilute phase at which the system reaches equilibrium. Once ϑ is given, the corresponding ϕ^* can be calculated. The remaining two equations are used to determine the actual volume when the dilute and droplet phases coexist, and these two equations enforce the conditions that the total molecular volume of the solute and solvent remain constant [70]. When the macroscopic threshold is exceeded, the protein switching reaction between the dilute and droplet phases dominates, and two distinct phases are macroscopically observed.

Under the condition that the number of proteins is fixed [consider only the switching of proteins between the dilute and droplet phases, corresponding to the two reactions in the second line of Eq. (1)], the steady-state probability of the dilute phase protein number n_+ in thermodynamic equilibrium obeys the Boltzmann-Gibbs distribution [71,72]:

$$P_{\mathrm{BG}}(N_+ = n_+ | N_{\mathrm{tot}} = n_{\mathrm{tot}}) = \frac{1}{Z_{n_{\mathrm{tot}}}} \exp\left[-\frac{F(n_+, n_-)}{k_B T}\right], \tag{5}$$

where $Z_{n_{\mathrm{tot}}} = \sum_{n_+=0}^{n_{\mathrm{tot}}} \exp[-F(n_+, n_-)/k_B T]$ is the partition function. At the equilibrium state, the rates $k_{\mathrm{in}}(n_+, n_-)$ and $k_{\mathrm{out}}(n_+, n_-)$ satisfy the detailed-balance condition:

$$\begin{aligned} &k_{\mathrm{out}}(n_+, n_-) P_{\mathrm{BG}}(N_+ = n_+ | N_{\mathrm{tot}} = n_{\mathrm{tot}}) \\ &\quad = k_{\mathrm{in}}(n_+ + 1, n_- - 1) P_{\mathrm{BG}}(N_+ = n_+ + 1 | N_{\mathrm{tot}} = n_{\mathrm{tot}}); \end{aligned}$$

then the relationship between $k_{\mathrm{in}}(n_+, n_-)$ and $k_{\mathrm{out}}(n_+, n_-)$ is described by the free energy as

$$\begin{aligned} k_{\mathrm{out}}(n_+, n_-) &= k_{\mathrm{in}}(n_+ + 1, n_- - 1) \\ &\quad \times \exp\left[\frac{F(n_+, n_-) - F(n_+ + 1, n_- - 1)}{k_B T}\right], \end{aligned} \tag{6}$$

and we assume that the rate $k_{\mathrm{in}}(n_+, n_-)$ is diffusion limited [41]:

$$k_{\mathrm{in}}(n_+, n_-) = \frac{6 D n_+}{(V_{\mathrm{tot}} - v n_-)^{2/3}}, \tag{7}$$

where D is the diffusion constant of the protein in the system, defined as $D = V_{\mathrm{tot}}^{2/3}/6\tau_D$, and τ_D is the average time of protein diffusion and can be used to measure the timescale of phase separation [41,73,74]. Hence, we can get $k_{\mathrm{out}}(n_+, n_-)$ through Eqs. (6) and (7). Here, $\omega = \tau_{\mathrm{D}}/\tau_{\mathrm{p}}$ is a scaling parameter that can be used to represent the difference between the time scales of phase separation and gene expression. Based on experimental observations, there is $\omega < 1$ [50,62–65,75].

C. Stochastic dynamics of phase-separating system with gene expression process

There is stochasticity in gene expression or cell-to-cell variation when only a single copy of the target gene and a few proteins are involved. The stochastic evolution of the gene expression process is usually described by a CME [24,76–83]. We established the CME of the reaction network Eq. (1). Let $P(n_+, n_-; t)$ be the joint probability distribution function of the system in the state where there are n_+ proteins in the dilute phase and n_- proteins in the droplet phase at time t. The CME describes the time evolution of probability:

$$\begin{aligned} \frac{\partial}{\partial t} P(n_+, n_-; t) &= k_{\mathrm{in}}(n_+ + 1, n_- - 1) P(n_+ + 1, n_- - 1; t) - k_{\mathrm{out}}(n_+, n_-) P(n_+, n_-; t) \\ &\quad + k_{\mathrm{out}}(n_+ - 1, n_- + 1) P(n_+ - 1, n_- + 1; t) - k_{\mathrm{in}}(n_+, n_-) P(n_+, n_-; t) \\ &\quad + \tilde{\lambda}(n_-) P(n_+ - 1, n_-; t) - \tilde{\lambda}(n_-) P(n_+, n_-; t) \\ &\quad + \delta[(n_+ + 1) P(n_+ + 1, n_-; t) - n_+ P(n_+, n_-; t)] + \delta[(n_- + 1) P(n_+, n_- + 1; t) - n_- P(n_+, n_-; t)], \end{aligned} \tag{8}$$

where the first two rows on the right-hand side describe the dynamics of phase separation, while the other terms correspond to the production and degradation of gene expression products.

The reaction rate equations can be used to analyze the macroscopic behavior of the corresponding system before we further perform stochastic analysis. Specifically, the rate equation corresponding to the constructed-above reaction network is

$$\begin{aligned} \frac{d\langle n_+ \rangle}{dt} &= \lambda + \frac{\zeta \langle n_- \rangle^h}{K^h + \langle n_- \rangle^h} + k_{\mathrm{out}}(\langle n_+ \rangle, \langle n_- \rangle) - k_{\mathrm{in}}(\langle n_+ \rangle, \langle n_- \rangle) - \delta \langle n_+ \rangle, \\ \frac{d\langle n_- \rangle}{dt} &= k_{\mathrm{in}}(\langle n_+ \rangle, \langle n_- \rangle) - k_{\mathrm{out}}(\langle n_+ \rangle, \langle n_- \rangle) - \delta \langle n_- \rangle. \end{aligned} \tag{9}$$

To investigate the regulatory role of phase separation processes in gene expression networks, we consider a system without phase separation compared with the model above, i.e., a conventional birth-and-death process with feedback. This stochastic model is described by the following two chemical reactions:

$$\begin{aligned} \mathrm{DNA} &\xrightarrow{\tilde{\lambda}(n_{\mathrm{tot}})} \mathrm{DNA} + N_{\mathrm{tot}}, \\ N_{\mathrm{tot}} &\xrightarrow{\delta} \emptyset. \end{aligned} \tag{10}$$

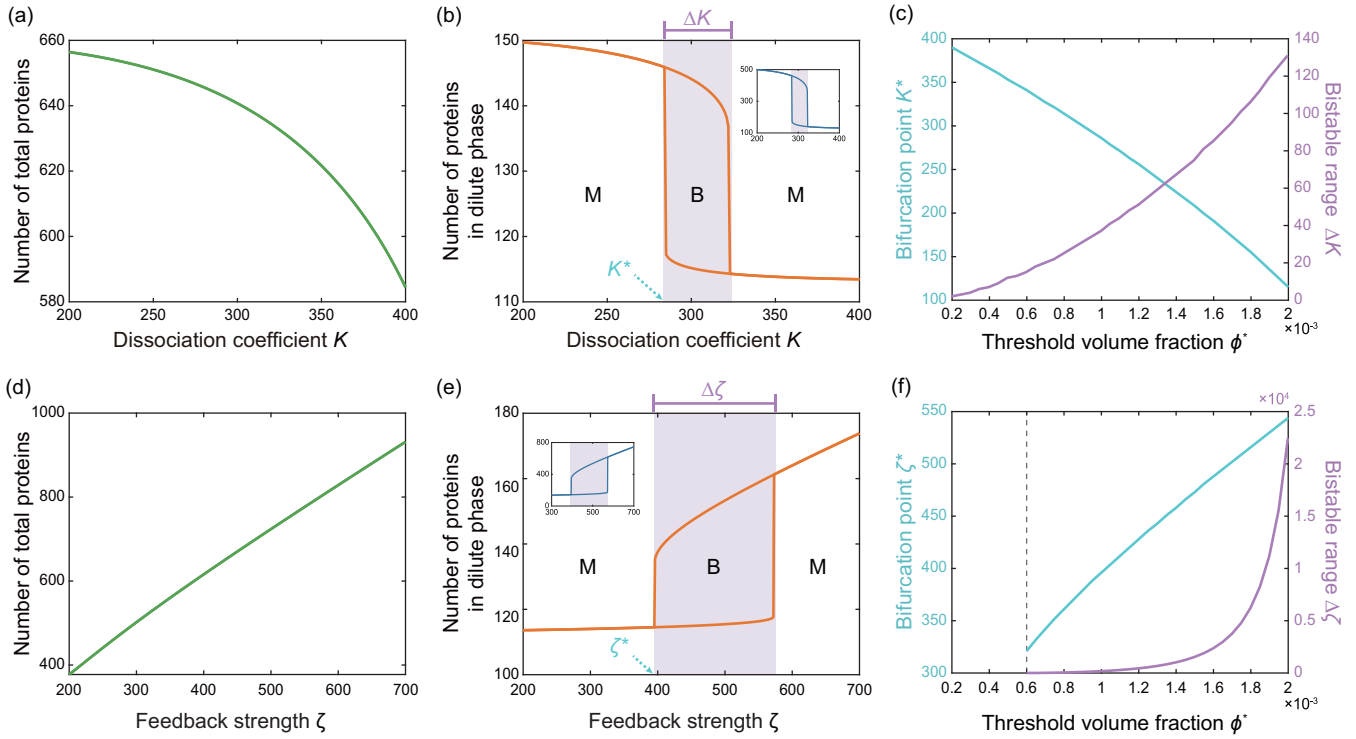

FIG. 2. Phase separation can induce bistable protein expression. (a)–(c) Effect of phase separation on the bifurcation behavior of K at $\lambda = 240$ and $K = 310$. (d)–(f) The role of phase separation in the bifurcation behavior of the feedback strength ζ at $\lambda = 240$ and $\zeta = 420$. (g)–(i) The effect of the phase separation on the bifurcation behavior of λ at $\zeta = 420$ and $K = 310$. (a), (d), and (h) The number of proteins at the stable point in a birth and death process with feedback. (b), (e), and (h) The bifurcation behavior of the phase separation system. The main graph shows the number of proteins in the dilute phase, and the subgraph shows the droplet phase of the corresponding system with the threshold volume fraction $\phi^* = 10^{-3}$. Purple areas indicate the parameter range of bistability (b); otherwise, it is monostable (m). (c), (f), and (i) The bifurcated parameter point (cyan lines) and the bistable region (purple lines) in the dilute phase as a function of the threshold volume fraction. The left side of the gray dotted line in (f) is monostable. The other parameter values are set as $\tau_p = 1$ and $\omega = 0.1$.

Unlike Eq. (2), the proteins are not partitioned, so the feedback is thought to be driven by the total protein copy number $n_{\rm tot}$:

$$\tilde{\lambda}(n_{\rm tot}) = \lambda + \frac{\zeta n_{\rm tot}^h}{K^h + n_{\rm tot}^h}. \tag{11}$$

Similarly, considering the stochastic fluctuations of proteins in Eq. (10), we obtain CME:

$$\frac{\partial P(n_{\rm tot};t)}{\partial t} = \tilde{\lambda}(n_{\rm tot}-1)P(n_{\rm tot}-1;t) - \tilde{\lambda}(n_{\rm tot})P(n_{\rm tot};t) + \delta[(n_{\rm tot}+1)P(n_{\rm tot}+1;t) - n_{\rm tot}P(n_{\rm tot};t)], \tag{12}$$

where $P(n_{\rm tot};t)$ is the probability distribution of the total number of proteins $n_{\rm tot}$ at t. The following reaction rate equation can describe the deterministic macroscopic dynamic properties of the system without phase separation:

$$\frac{d\langle n_{\rm tot}\rangle}{dt} = \lambda + \frac{\zeta \langle n_{\rm tot}\rangle^h}{K^h + \langle n_{\rm tot}\rangle^h} - \delta\langle n_{\rm tot}\rangle. \tag{13}$$

The rate equations Eqs. (9) and (13) do not include the inherent fluctuation of the birth-death process. Note that, if $k_{\rm in}(n_+, n_-) = k_{\rm out}(n_+, n_-) = 0$, the system described by Eq. (1) degenerates into the birth-death process with feedback Eq. (10).

III. RESULTS

A. Phase separation can induce bistability

We numerically investigate the impact of phase separation on the macroscopic behavior of gene expression products. Numerical results for the steady state of Eqs. (9) and (13) are shown in Figs. 2–6.

The bifurcation diagram illustrates that the system has bistable properties under different conditions. To investigate the effect of phase separation on the dynamic properties of the system, we analyze the bifurcation behavior of feedback parameters, including the dissociation coefficient K, the basal generation rate λ, and the feedback strength ζ, in the conventional and phase-separated birth-and-death models, respectively. The bifurcation diagrams demonstrate the coexistence range of two states and the transition between the two stable states in the entire system, dilute phase, and droplet phase. The mean steady-state protein levels in the dilute and droplet phases change similarly to conventional birth-and-death processes with positive feedback. However, phase separation can induce bistability in the parameter range where the gene expression product is monostable in the conventional birth-and-death process with feedback [Figs. 2(a) and 2(b), 2(d) and 2(e), and 3(a) and 3(d)]. The bistability of the droplet phase is consistent with the dilute phase, while the two stable states in the droplet phase have a large gap [insets in

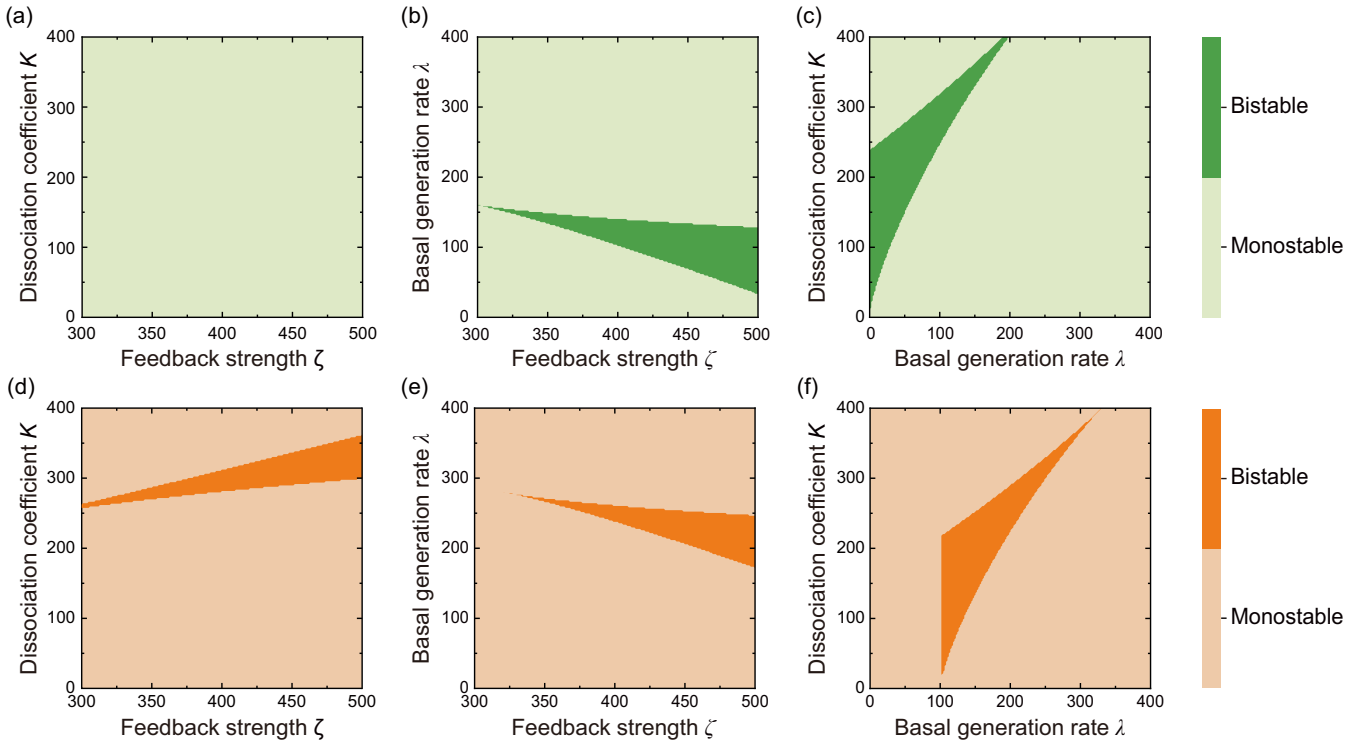

FIG. 3. Heatmaps of the number of stable points on the feedback parameter plane. The number of stable points is given as a binary function of $K-\zeta$ at $\lambda = 240$ (left), $\lambda-\zeta$ at $K = 310$ (middle), and $K-\lambda$ at $\zeta = 420$ (right). (a)–(c) Bifurcating behavior of the conventional birth and death process with positive feedback. (d)–(f) The number of stable points for the system with phase separation. The dark and light regions correspond to bistable and monostable, respectively. The other parameter value is set as $\tau_p = 1$ and $\omega = 0.1$.

Figs. 2(b), 2(e), and 2(h)]. This result suggests that the droplet phase plays a significant role in phenotype switching.

We investigate the specific impact of phase separation on the bistability of feedback parameters. Figures 2(c) and 2(f) demonstrate that phase separation can control the bistable characteristics of the system, including the bifurcation point (ζ^* and K^*) and bistable range ($\Delta\zeta$ and ΔK). As the threshold volume fraction ϕ^* increases, $\Delta\zeta$ and ΔK expand, which is significant for $\Delta\zeta$. This suggests that LLPS is closely related to feedback strength. However, it is observed that ϕ^* does not affect the bistable range of λ [Fig. 2(i)]. Figure 4 shows heatmaps for the number of steady states in the parameter space $K-\zeta$ at different ϕ^*. An increase in ϕ^* causes the bistable region to shift and expand. The bistability region extends, meaning more parameter cases can form bistability.

The threshold volume fraction ϕ^* describes the level of difficulty at which phase separation occurs. We found that the number of proteins in the dilute and droplet phases produced a bistable behavior with respect to ϕ^* [Figs. 5(a) and 5(b)]. The number of proteins in the dilute phase increases with increasing threshold, and the opposite is true for droplets. In the bistable region, the two stable states are large droplets with

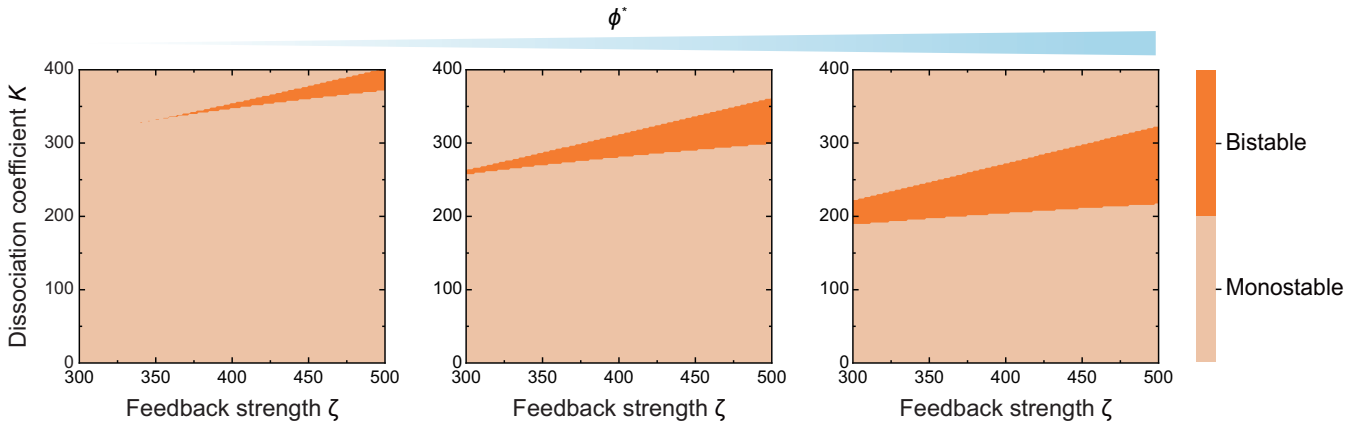

FIG. 4. Heatmaps of stable points with different threshold volume fractions. ϕ^* is 0.5×10^{-3}, 10^{-3}, and 1.5×10^{-3}, respectively. The dark and light regions correspond to bistable and monostable. The other parameter values are set as $\lambda = 240$, $\tau_p = 1$, and $\omega = 0.1$.

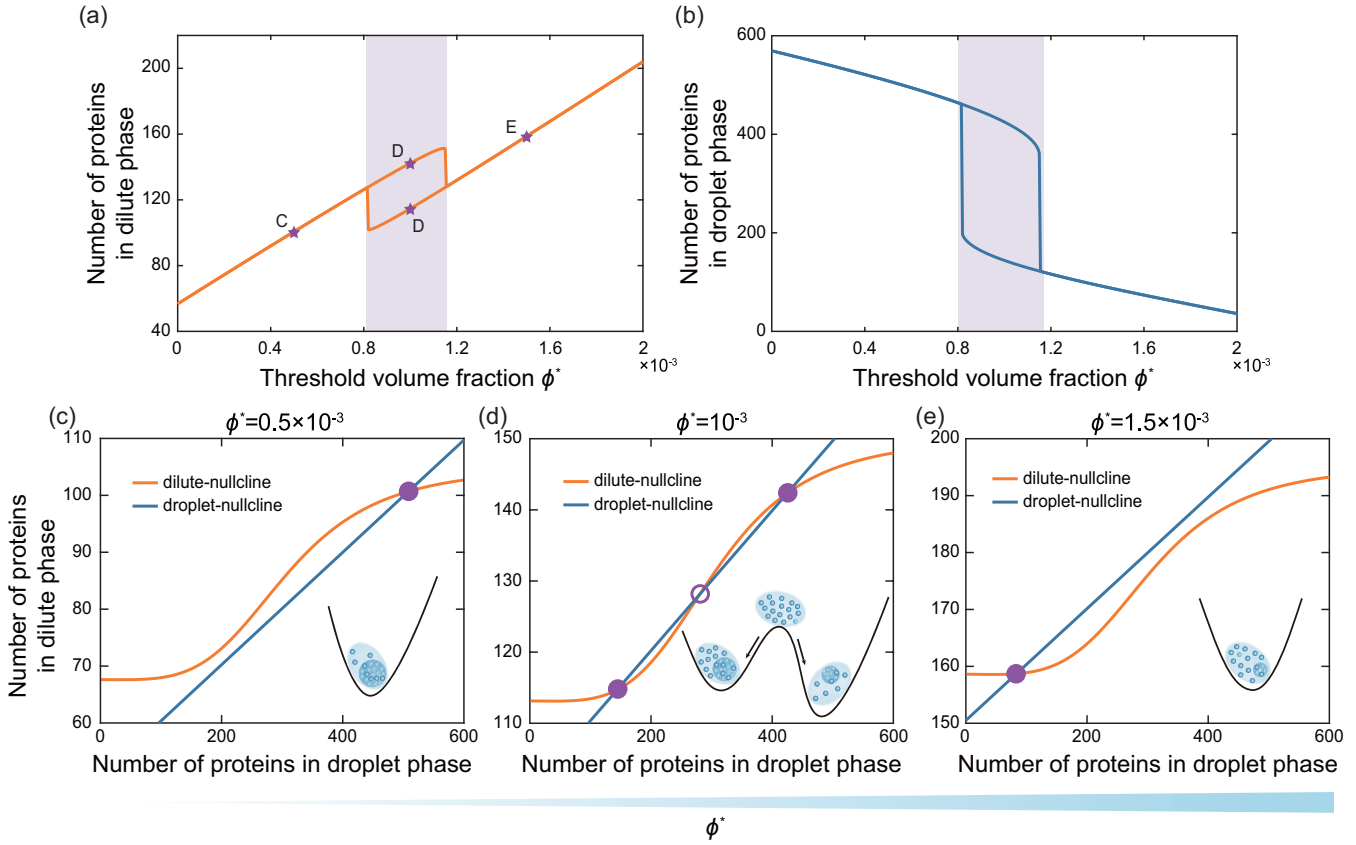

FIG. 5. The number of proteins in the dilute and droplet phases exhibit bistability for the threshold volume fraction. (a) The number of proteins at stable points in the dilute phase as a function of ϕ^*. (b) Bifurcation behavior of proteins in droplet phase with respect to ϕ^*. (c)–(e) Nullclines for protein in the dilute phase and droplet phase, corresponding to the stars in (a). The subplots show a schematic of the system in a steady state. The other parameter values are set as $\tau_p = 1$, $\omega = 0.1$, $\lambda = 240$, $K = 310$, and $\zeta = 420$.

more protein in the dilute phase and small droplets with less protein in the dilute phase. Figures 5(c)–5(e) show the dilute-nullcline and the droplet-nullcline at different ϕ^*, where the number of crossings defines the number of fixed points within the deterministic system. There is only one stable steady state when ϕ^* is smaller [Fig. 5(c)] or larger [Fig. 5(e)]. Three fixed points appear in the bistable region, of which two stable fixed points are separated by an unstable one [Fig. 5(d)]. This demonstrates that the phase separation regulation is capable of bistability in the dilute and droplet phases with consistent kinetics. When $\phi^* > 2.4 \times 10^{-3}$, the system Eq. (1) has no fixed point and degenerates to Eq. (10).

B. Phase separation can induce bimodal distribution

In the previous section, we analyzed the macroscopic behavior of phase separation in deterministic systems. In this section, we will focus on the effects of phase separation on stochastic system dynamics. The CME provides the most complete information about the probabilistic behavior of biochemical reaction networks. However, solving the master Eqs. (12) and (8) is a difficult task due to the nonlinear terms arising from the propensity functions $k_{\text{in}}(n_+, n_-)$ and $k_{\text{out}}(n_+, n_-)$ involving the phase separation process. We obtain the results by simulation using the Gillespie algorithm [78].

We observe from the time series that the gene expression process with positive feedback has no obvious transition behavior, and only noise-induced fluctuations exist in Fig. 6(a). However, two switching states are clearly visible in the dilute and droplet phases [Fig. 6(b)]. Figure 6(c) corresponds to the probability of stationary total protein distributions in Fig. 6(a), which is a unimodal distribution. We find that the stationary joint distribution of proteins for the dilute and droplet phases is bimodal in Fig. 6(d), which means that the system is capable of generating peaks with different probabilities for multivariate. One corresponds to small droplets with low total protein and the other to large droplets with high total protein. The marginal distributions in the dilute and droplet phases showed a bimodal distribution. In addition, the time evolution series in Fig. 6(c) shows a bimodal distribution of the sum of proteins in the dilute and droplet phases. Therefore, we conclude that phase separation is a mechanism inducing bimodality. If both dilute and droplet phases exhibit bimodality, the bimodality of the latter is more pronounced than that of the former, showing obvious high and low expression levels. Bimodal distributions are commonly associated with cell fate differentiation and determination, the basis of organism development and mis-development.

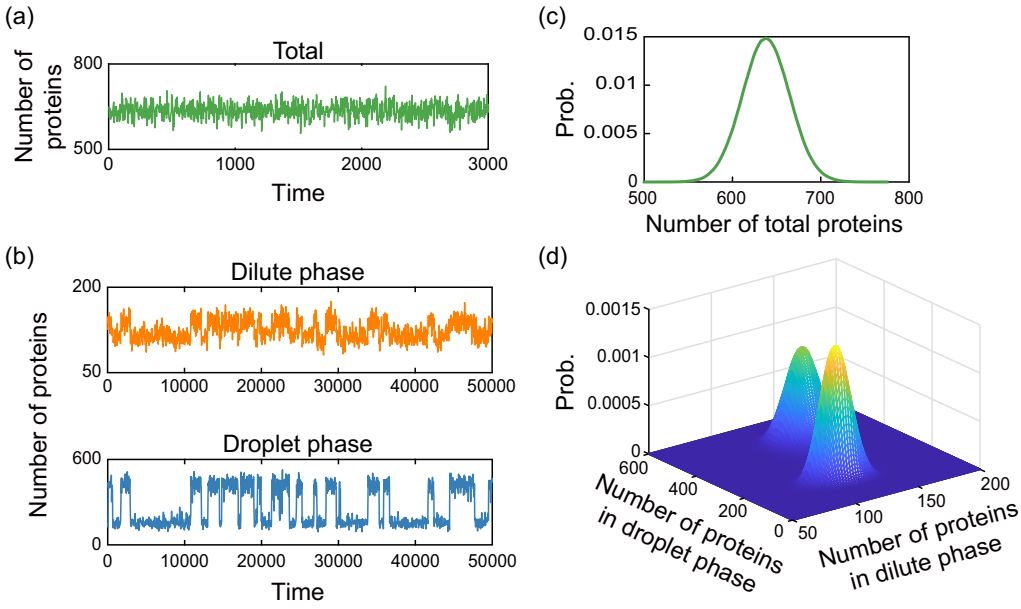

FIG. 6. Phase separation can induce bimodal protein expression in the presence of feedback. (a) The stochastic trajectory of the birth and death model with positive feedback. (b) The evolution trajectories of the protein in the dilute phase (top) and the droplet phase (below). (c) The histogram shows a steady-state distribution of protein corresponding to (a). (d) The joint distribution of proteins in the dilute phase and droplet phase, which corresponds to (b). The other parameter values are set as $\tau_p = 1$, $\omega = 0.1$, $\lambda = 240$, $\zeta = 420$, and $K = 310$.

C. Differences in bimodality between dilute and droplet phases

Next, we further quantify the bimodality by the distance between double peak points of the probability and the probability difference between these peak points.

Like the deterministic system, the protein numbers in the dilute and droplet phases show a bimodal region for the threshold volume fraction ϕ^* [Figs. 7(a) and 7(b)]. The bimodal region in the dilute phase is much narrower than in the droplet phase. The bimodality of the dilute and droplet phases is inconsistent due to stochastic dynamics. Further, the distribution can appear in five modes as ϕ^* increases. Figure 7(c) shows the evolutionary process of peak numbers and peak probabilities with ϕ^*, and experiments have shown that, under certain conditions, proteins can form droplets of different sizes [84,85]. In the droplet phase, the peak far away from the origin goes under with amplifying ϕ^*, while the peak close to the origin begins to grow, and a bimodal then occurs. When the heights of the two peaks are roughly equal in the droplet phase, the protein distribution in the dilute phase exhibits a bimodal pattern. As ϕ^* continues to grow, the bimodal distribution in the dilute and droplet phases disappears sequentially. The results indicate that a droplet plays a significant role in regulating the relative switching of phenotypes. In addition, the doublet peaks show a high dependence on the feedback parameters (Fig. 8).

IV. CONCLUSIONS

Bistability and even multistability play an essential role in cell differentiation. The development of an organism can be viewed as a sequence of transitions between different cell fates. Authors of previous studies have suggested that protein bistability and bimodality are associated with cell fate decisions and differentiation, a mechanism that contributes to the phenotypic diversity for cellular adaptation and growth in fluctuating environments [86]. LLPS is widely present in gene expression, but authors of previous studies of phenotypic switching have not considered its role in positive feedback. Therefore, effective models are needed to study the impact of LLPS on gene expression products with positive feedback. Here, we propose a multiscale model that couples the LLPS and gene expression processes to capture the distribution of proteins induced by phase separation in autoregulation dynamics. The LLPS process has a relatively fast time scale, while the gene expression process has a relatively slow time scale.

First, we employ thermodynamic theories to determine a threshold value that reveals two distinct phases at a macroscopic level. Next, we describe the evolutionary dynamics of proteins between the dilute and droplet phases through stochastic exchange events and determine the trend function for phase separation at thermodynamic equilibrium. The range of bistability in feedback strength and dissociation coefficient expands with an increase in the phase separation threshold volume fraction, when the basic generation rate stays constant. This observation implies that phase separation could serve as a potential mechanism for regulating hysteresis loops [87]. Future studies will further investigate the effect of phase separation on hysteresis. Finally, phase separation can induce bimodal distributions in stochastic systems. This double peak is revealed in three aspects: the joint distribution, the marginal distribution, and the distribution of the sum of proteins in the dilution and droplet phases. Stochastic dynamics causes the double peaks in the dilute and droplet phases to be asynchronous. When the droplet phase produces two peaks with

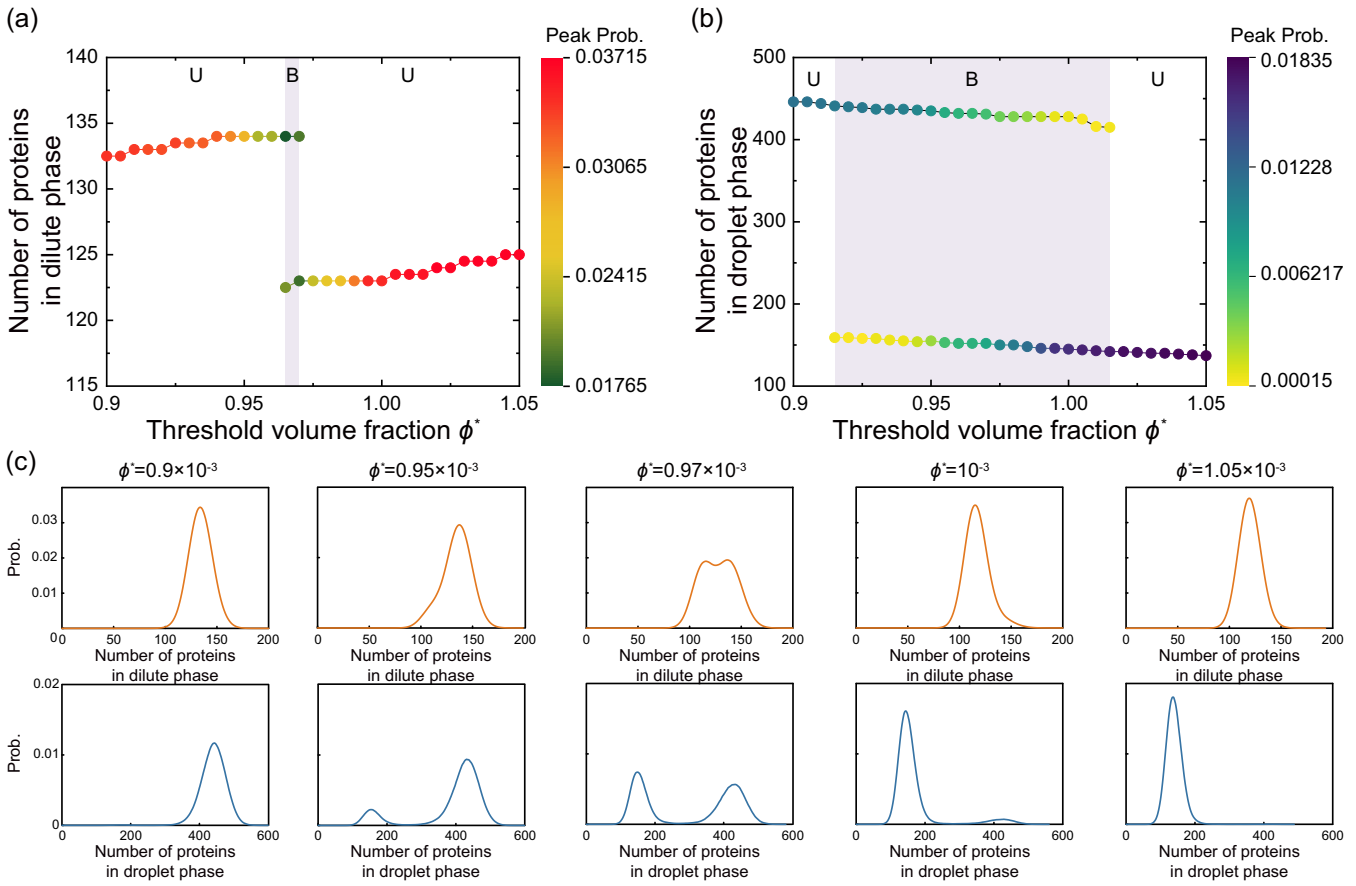

FIG. 7. Phase separation can induce bimodal protein expression in the presence of feedback. (a) The most probable protein numbers in the dilute phase as a function of ϕ^*. (b) The most probable protein numbers in the droplet phase as a function of ϕ^*. The color bars represent the peak probability. The purple area is bimodal (b); otherwise, it is unimodal (U). (c) The example of unimodal/bimodal protein distributions. The first row is the dilute phase, and the second row is the droplet phase. The histograms are generated using data from 10^4 independent Gillespie simulations, each lasting 3×10^4 s. The other parameter values are set $\tau_p = 1$, $\omega = 0.1$, $\lambda = 240$, $K = 310$, and $\zeta = 420$.

equivalent levels, it can result in double peaks in the dilute phase. This discovery provides insights into explaining the phenomenon of the droplet phase buffering the dilute phase.

Theoretical models describing critical phenomena have been studied [88]. In our analysis, we considered two factors: the total volume and the total number of proteins, with the total volume remaining constant throughout. In Sec. II B, a constant protein number is maintained due to the rapid phase separation compared with gene expression, which classifies the system as model B. In Sec. II A, the gene expression process results in a change in the total number of proteins, which is the interaction between the nonconserved field (protein number) and the conserved field (volume), so the system is model C. In addition, when the cell cycle is considered, both the total number of proteins and the total volume can change, which can be classified as model A. If more complex conditions are considered, it may correspond to other models.

We employed the Hill function to model autoregulation, noting its limitations, particularly its potential inaccuracies in scenarios where basal production rates are low and gene-protein binding rates are significantly faster than unbinding rates [89]. Future researchers should explore the role of phase separation within a broader range of autoregulatory models to enhance our understanding comprehensively.

The modeling method used in this paper is universal. Our research can be extended to more complex biochemical processes, such as the two-state model describing transcriptional bursting [45,90]. Promoter leakage can also be considered, as leakage rates are strongly correlated with bistability [91]. Furthermore, the model in this paper is based on the Markov hypothesis, which states that the stochastic dynamics of the reactants are only affected by the current state of the system. Authors of more studies in recent years have shown that non-Markovian models of biochemical dynamics can effectively simulate many intermediate biochemical processes [60,92]. Therefore, we attempt to analyze the impact of LLPS on the kinetic properties of gene expression models with arbitrary waiting time distributions.

In addition, the cellular interior is highly crowded with numerous protein species, which presents the problem of multicomponent phase separation with the need to account for the different interactions between species [70,93]. In recent years, remarkable results have been achieved in studying the phase behavior of multicomponent systems. Physically driven methods can simulate systems with more than a single component, while theoretical methods can simulate the factors of multivalent and cooperative interactions that lead to phase separation. On the other hand, the inherent complexity of

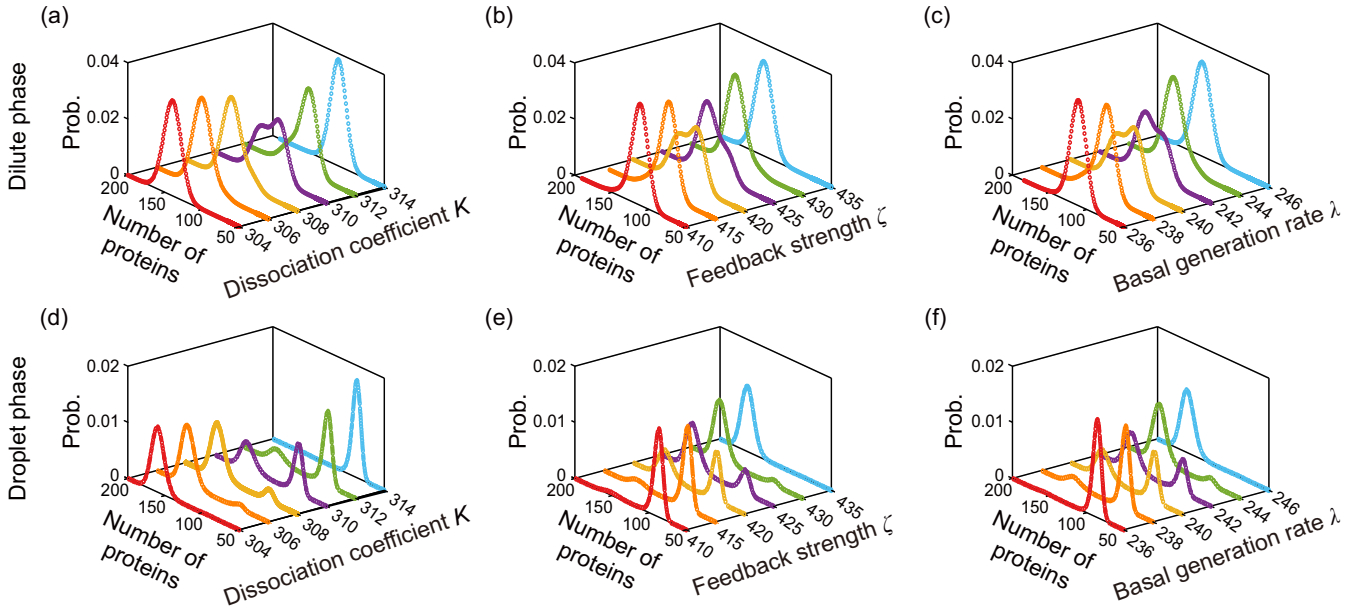

FIG. 8. Stationary protein distributions in the dilute phase and droplet phase. (a)–(c) Protein distribution in dilute phase. (d)–(f) Protein distribution in droplet phase. (a) and (d) The probability distribution under different K, where $\lambda = 240$ and $\zeta = 420$. (b) and (e) The probability distribution under different ζ, where $\lambda = 240$ and $K = 310$. (c) and (f) The probability distribution under different λ, where $\zeta = 420$ and $K = 310$. The histograms are generated using data from 10^4 independent Gillespie simulations, each lasting 3×10^4 s. The other parameter values are set $\tau_p = 1$ and $\omega = 0.1$.

biological systems is often described by high-dimensional dynamical models. The dimension reduction approach of landscape offers an effective research tool to investigate the underlying structural characteristics of the dynamical systems [94]. Adding phase separation to these models can provide insights into the regulation of gene expression, such as analyzing how phase separation affects the tristability or intermediate states of the system. However, it is generally recognized that the most critical milestone at the interface of computational science and molecular biology is the application of deep learning of evolutionary information [95,96], and we hope to solve the modeling problem of multiphase systems through machine-learning methods. In general, future work will be needed to elucidate better how the environment within crowded cells modulates phenotypic properties.

Finally, our research can serve as a foundation for research in fields such as synthetic biology. Our results show that phase separation can effectively regulate the diversity of biological phenotypes and provide guidance for biologists to design synthetic circuits to explore how phase separation regulates bistability and phenotypes.

ACKNOWLEDGMENTS

This paper was supported by Natural Science Foundation of China Grants No. 12171494 and No. 12301646; by Key-Area Research and Development Program of Guangzhou, China Grants No. 202007030004 and No. 2019B110233002; by Guangdong Basic and Applied Basic Research Foundation Grants No. 2022A1515011540, No. 2024A1515012786, No. 2023A1515110273, and No. 2023A1515011982; by Fundamental Research Funds for the Central Universities, Sun Yat-sen University Grants No. 23qnpy48 and No. 23qnpy49; and by China Postdoctoral Science Foundation Grant No. 2023M734061.

[1] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, A stochastic single-molecule event triggers phenotype switching of a bacterial cell, Science **322**, 442 (2008).

[2] M. Acar, A. Becskei, and A. Van Oudenaarden, Enhancement of cellular memory by reducing stochastic transitions, Nature (London) **435**, 228 (2005).

[3] P. B. Gupta, C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser, and E. S. Lander, Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells, Cell **146**, 633 (2011).

[4] T. S. Gardner, C. R. Cantor, and J. J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, Nature (London) **403**, 339 (2000).

[5] H. Kobayashi, M. Kaern, M. Araki, K. Chung, T. S. Gardner, C. R. Cantor, and J. J. Collins, Programmable cells: Interfacing natural and engineered gene networks, Proc. Natl. Acad. Sci. USA **101**, 8414 (2004).

[6] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. Van Oudenaarden, Multistability in the lactose utilization network of *Escherichia coli*, Nature (London) **427**, 737 (2004).

[7] W. Xiong and J. E. Ferrell Jr, A positive-feedback-based bistable 'memory module' that governs a cell fate decision, Nature (London) **426**, 460 (2003).

[8] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins, Noise-based switches and amplifiers for gene expression, Proc. Natl. Acad. Sci. USA **97**, 2075 (2000).

[9] J. Paulsson, Summing up the noise in gene networks, Nature (London) **427**, 415 (2004).

[10] X. Yang, S. Luo, Z. Zhang, Z. Wang, T. Zhou, and J. Zhang, Silent transcription intervals and translational bursting lead to diverse phenotypic switching, Phys. Chem. Chem. Phys. **24**, 26600 (2022).

[11] H. Maamar and D. Dubnau, Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop, Mol. Microbiol. **56**, 615 (2005).

[12] P. Gamba, M. J. Jonker, and L. W. Hamoen, A novel feedback loop that controls bimodal expression of genetic competence, PLoS Genet., **11**, e1005047 (2015).

[13] D. Fraser and M. Kaern, A chance at survival: Gene expression noise and phenotypic diversification strategies, Mol. Microbiol. **71**, 1333 (2009).

[14] E. Kussell and S. Leibler, Phenotypic diversity, population growth, and information in fluctuating environments, Science **309**, 2075 (2005).

[15] M. Thattai and A. Van Oudenaarden, Stochastic gene expression in fluctuating environments, Genetics **167**, 523 (2004).

[16] A. Becskei and L. Serrano, Engineering stability in gene networks by autoregulation, Nature (London) **405**, 590 (2000).

[17] M. A. Savageau, Comparison of classical and autogenous systems of regulation in inducible operons, Nature (London) **252**, 546 (1974).

[18] S. Maloy and V. Stewart, Autogenous regulation of gene expression, J. Bacteriol. **175**, 307 (1993).

[19] M. A. Savageau, Significance of autogenously regulated and constitutive synthesis of regulatory proteins in repressible biosynthetic systems, Nature (London) **258**, 208 (1975).

[20] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network motifs: Simple building blocks of complex networks, Science **298**, 824 (2002).

[21] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, Network motifs in the transcriptional regulation network of *Escherichia coli*, Nat. Genet. **31**, 64 (2002).

[22] J. Holehouse, Z. Cao, and R. J. B. J. Grima, Stochastic modeling of autoregulatory genetic feedback loops: A review and comparative study, Biophys. J. **118**, 1517 (2020).

[23] D. W. Austin, M. S. Allen, J. M. McCollum, R. D. Dar, J. R. Wilgus, G. S. Sayler, N. F. Samatova, C. D. Cox, and M. L. Simpson, Gene network shaping of inherent noise spectra, Nature (London) **439**, 608 (2006).

[24] M. Thattai and A. Van Oudenaarden, Intrinsic noise in gene regulatory networks, Proc. Natl. Acad. Sci. USA **98**, 8614 (2001).

[25] Y. Dublanche, K. Michalodimitrakis, N. Kummerer, M. Foglierini, and L. Serrano, Noise in transcription negative feedback loops: Simulation and experimental analysis, Mol. Syst. Biol. **2**, 41 (2006).

[26] M. E. Wall, W. S. Hlavacek, and M. A. Savageau, Design of gene circuits: Lessons from bacteria, Nat. Rev. Genet. **5**, 34 (2004).

[27] N. Rosenfeld, M. B. Elowitz, and U. Alon, Negative autoregulation speeds the response times of transcription networks, J. Mol. Biol. **323**, 785 (2002).

[28] A. K. Shalek, R. Satija, X. Adiconis, R. S. Gertner, J. T. Gaublomme, R. Raychowdhury, S. Schwartz, N. Yosef, C. Malboeuf, D. Lu *et al.*, Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells, Nature (London) **498**, 236 (2013).

[29] A. Raj, S. A. Rifkin, E. Andersen, and A. Van Oudenaarden, Variability in gene expression underlies incomplete penetrance, Nature (London) **463**, 913 (2010).

[30] G. M. Süel, J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz, An excitable gene regulatory circuit induces transient cellular differentiation, Nature (London) **440**, 545 (2006).

[31] N. Barkai and S. Leibler, Robustness in simple biochemical networks, Nature (London) **387**, 913 (1997).

[32] T. Hu, H. Yamagishi, J. Maeda, J. McAnally, C. Yamagishi, and D. Srivastava, Tbx1 regulates fibroblast growth factors in the anterior heart field through a reinforcing autoregulatory loop involving forkhead transcription factors, Development **131**, 5491 (2004).

[33] D. T. Odom, N. Zizlsperger, D. B. Gordon, G. W. Bell, N. J. Rinaldi, H. L. Murray, T. L. Volkert, J. Schreiber, P. A. Rolfe, D. K. Gifford *et al.*, Control of pancreas and liver gene expression by HNF transcription factors, Science **303**, 1378 (2004).

[34] J. E. Pimanda, I. J. Donaldson, M. F. de Bruijn, S. Kinston, K. Knezevic, L. Huckle, S. Piltz, J. R. Landry, A. R. Green, D. Tannahill *et al.*, The SCL transcriptional network and BMP signaling pathway interact to regulate RUNX1 activity, Proc. Natl. Acad. Sci. USA **104**, 840 (2007).

[35] L. T. Grajkowska, M. Ceribelli, C. M. Lau, M. E. Warren, I. Tiniakou, S. Nakandakari Higa, A. Bunin, H. Haecker, L. A. Mirny, L. M. Staudt *et al.*, Isoform-specific expression and feedback regulation of E protein TCF4 control dendritic cell lineage specification, Immunity **46**, 65 (2017).

[36] K. Srinivasan, D. P. Leone, R. K. Bateson, G. Dobreva, Y. Kohwi, T. Kohwi-Shigematsu, R. Grosschedl, and S. K. McConnell, A network of genetic repression and derepression specifies projection fates in the developing neocortex, Proc. Natl. Acad. Sci. USA **109**, 19071 (2012).

[37] K. Oishi, M. Aramaki, and K. Nakajima, Mutually repressive interaction between Brn1/2 and Rorb contributes to the establishment of neocortical layer 2/3 and layer 4, Proc. Natl. Acad. Sci. USA **113**, 3371 (2016).

[38] H. A. Harrington, E. Feliu, C. Wiuf, and M. P. Stumpf, Cellular compartments cause multistability and allow cells to process more information, Biophys. J. **104**, 1824 (2013).

[39] D. Tenenbaum, J. I. Marrone, H. E. Grecco, and A. C. Ventura, Robustness in spatially driven bistability in signaling systems, Sci. Rep. **10**, 5591 (2020).

[40] O. S. Venturelli, H. El-Samad, and R. M. Murray, Synergistic dual positive feedback loops established by molecular sequestration generate robust bimodal response, Proc. Natl. Acad. Sci. USA **109**, E3324 (2012).

[41] A. Klosin, F. Oltsch, T. Harmon, A. Honigmann, F. Julicher, A. A. Hyman, and C. Zechner, Phase separation provides a mechanism to reduce noise in cells, Science **367**, 464 (2020).

[42] S. Alberti, A. Gladfelter, and T. Mittag, Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates, Cell **176**, 419 (2019).

[43] E. E. Griffin, D. J. Odde, and G. Seydoux, Regulation of the MEX-5 gradient by a spatially segregated kinase/phosphatase cycle, Cell **146**, 955 (2011).

[44] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry, Nat. Rev. Mol. Cell Biol. **18**, 285 (2017).

[45] L. Hong, Z. Wang, Z. Zhang, S. Luo, T. Zhou, and J. Zhang, Phase separation reduces cell-to-cell variability of transcriptional bursting, Math. Biosci. **367**, 109127 (2024).
[46] B. Y. Zaslavsky and V. N. Uversky, *In aqua veritas*: The indispensable yet mostly ignored role of water in phase separation and membrane-less organelles, Biochemistry **57**, 2437 (2018).
[47] C. A. Azaldegui, A. G. Vecchiarelli, and J. S. Biteen, The emergence of phase separation as an organizing principle in bacteria, Biophys. J. **120**, 1123 (2021).
[48] C. D. Crowe and C. D. Keating, Liquid-liquid phase separation in artificial cells, Interface Focus **8**, 20180032 (2018).
[49] A.-M. Ladouceur, B. S. Parmar, S. Biedzinski, J. Wall, S. G. Tope, D. Cohn, A. Kim, N. Soubry, R. Reyes-Lamothe, and S. C. Weber, Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid-liquid phase separation, Proc. Natl. Acad. Sci. USA **117**, 18540 (2020).
[50] A. A. Hyman, C. A. Weber, and F. Julicher, Liquid-liquid phase separation in biology, Annu. Rev. Cell Dev. Biol. **30**, 39 (2014).
[51] A. F. Harrison and J. Shorter, RNA-binding proteins with prion-like domains in health and disease, Biochem. J. **474**, 1417 (2017).
[52] L. Peng, E. M. Li, and L. Y. Xu, From start to end: Phase separation and transcriptional regulation, BBA-Gene Regul. Mech. **1863**, 194641 (2020).
[53] A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, and N. M. Hannett, Transcription factors activate genes through the phase-separation capacity of their activation domains, Cell **175**, 1842 (2018).
[54] D. Cai, D. Feliciano, P. Dong, E. Flores, M. Gruebele, N. Porat-Shliom, S. Sukenik, Z. Liu, and J. Lippincott-Schwartz, Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression, Nat. Cell Biol. **21**, 1578 (2019).
[55] S. Mehta and J. Zhang, Liquid-liquid phase separation drives cellular function and dysfunction in cancer, Nat. Rev. Cancer **22**, 239 (2022).
[56] Z. Wang, S. Luo, Z. Zhang, T. Zhou, and J. Zhang, 4D nucleome equation predicts gene expression controlled by long-range enhancer-promoter interaction, PLoS Comput. Biol. **19**, e1011722 (2023).
[57] B. R. Sabari, A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, and J. C. Manteiga, Coactivator condensation at super-enhancers links phase separation and gene control, Science **361**, eaar3958 (2018).
[58] U. Alon, Network motifs: Theory and experimental approaches, Nat. Rev. Genet. **8**, 450 (2007).
[59] J. Frank and C. M. Spahn, The ribosome and the mechanism of protein synthesis, Rep. Prog. Phys. **69**, 1383 (2006).
[60] Z. Wang, Z. Zhang, and T. Zhou, Analytical results for non-Markovian models of bursty gene expression, Phys. Rev. E **101**, 052406 (2020).
[61] J. Zhang and T. Zhou, Markovian approaches to modeling intracellular reaction processes with molecular memory, Proc. Natl. Acad. Sci. USA **116**, 23542 (2019).
[62] A. J. Larsson, P. Johnsson, M. Hagemann-Jensen, L. Hartmanis, O. R. Faridani, B. Reinius, Å. Segerstolpe, C. M. Rivera, B. Ren, and R. Sandberg, Genomic encoding of transcriptional burst kinetics, Nature (London) **565**, 251 (2019).
[63] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Germline P granules are liquid droplets that localize by controlled dissolution/condensation, Science **324**, 1729 (2009).
[64] X. Su, J. A. Ditlev, E. Hui, W. Xing, S. Banjade, J. Okrut, D. S. King, J. Taunton, M. K. Rosen, and R. D. Vale, Phase separation of signaling molecules promotes T cell receptor signal transduction, Science **352**, 595 (2016).
[65] A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, and T. M. Franzmann, A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation, Cell **162**, 1066 (2015).
[66] P. Thomas, N. Popović, and R. Grima, Phenotypic switching in gene regulatory networks, Proc. Natl. Acad. Sci. USA **111**, 6994 (2014).
[67] N. Friedman, L. Cai, and X. S. Xie, Linking stochastic dynamics to population distribution: An analytical framework of gene expression, Phys. Rev. Lett. **97**, 168302 (2006).
[68] C. A. Weber, D. Zwicker, F. Jülicher, and C. F. Lee, Physics of active emulsions, Rep. Prog. Phys. **82**, 064601 (2019).
[69] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis: Zweiter Band: Funktionentheorie · Nullstellen Polynome · Determinanten Zahlentheorie* (Springer, Berlin, 1925).
[70] D. Deviri and S. A. Safran, Physical theory of biological noise buffering by multicomponent phase separation, Proc. Natl. Acad. Sci. USA **118**, e2100099118 (2021).
[71] A. Carati, Thermodynamics and time averages, Physica A **348**, 110 (2005).
[72] C. Tsallis, Remarks on the nonuniversality of Boltzmann-Gibbs statistical mechanics, Fractals **11**, 319 (2003).
[73] R. Yamamoto and A. Onuki, Dynamics of highly supercooled liquids: Heterogeneity, rheology, and diffusion, Phys. Rev. E **58**, 3515 (1998).
[74] A. Ochab-Marcinek and R. Hołyst, Scale-dependent diffusion of spheres in solutions of flexible and rigid polymers: Mean square displacement and autocorrelation function for FCS and DLS measurements, Soft Matter **7**, 7366 (2011).
[75] A. R. Dörrbaum, L. Kochen, J. D. Langer, and E. M. Schuman, Local and global influences on protein turnover in neurons and glia, eLife **7**, e34202 (2018).
[76] H. H. McAdams and A. Arkin, Stochastic mechanisms in gene expression, Proc. Natl. Acad. Sci. USA **94**, 814 (1997).
[77] J. Paulsson and M. Ehrenberg, Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks, Phys. Rev. Lett. **84**, 5447 (2000).
[78] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. C **81**, 2340 (1977).
[79] T. B. Kepler and T. C. Elston, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations, Biophys. J. **81**, 3116 (2001).
[80] J. M. Raser and E. K. O'Shea, Control of stochasticity in eukaryotic gene expression, Science **304**, 1811 (2004).
[81] P. S. Gutierrez, D. Monteoliva, and L. Diambra, Role of cooperative binding on noise expression, Phys. Rev. E **80**, 011914 (2009).
[82] Á. Sánchez and J. Kondev, Transcriptional control of noise in gene expression, Proc. Natl. Acad. Sci. USA **105**, 5081 (2008).
[83] M. Chen, S. Luo, M. Cao, C. Guo, T. Zhou, and J. Zhang, Exact distributions for stochastic gene expression models with

arbitrary promoter architecture and translational bursting, Phys. Rev. E **105**, 014405 (2022).

[84] E. W. Martin, T. S. Harmon, J. B. Hopkins, S. Chakravarthy, J. J. Incicco, P. Schuck, A. Soranno, and T. Mittag, A multi-step nucleation process determines the kinetics of prion-like domain phase separation, Nat. Commun. **12**, 4513 (2021).

[85] M. Kar, F. Dar, T. J. Welsh, L. T. Vogel, R. Kühnemuth, A. Majumdar, G. Krainer, T. M. Franzmann, S. Alberti, and C. A. Seidel, Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions, Proc. Natl. Acad. Sci. USA **119**, e2202222119 (2022).

[86] D. Arendt, J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler *et al.*, The origin and evolution of cell types, Nat. Rev. Genet. **17**, 744 (2016).

[87] M. Pájaro, I. Otero-Muras, C. Vázquez, and A. A. Alonso, Transient hysteresis and inherent stochasticity in gene regulatory networks, Nat. Commun. **10**, 4581 (2019).

[88] U. C. Täuber, *Critical Dynamics: A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior* (Cambridge University Press, Cambridge, 2014).

[89] J. Holehouse and R. Grima, Revisiting the reduction of stochastic models of genetic feedback loops with fast promoter switching, Biophys. J. **117**, 1311 (2019).

[90] Z. Wang, Z. Zhang, S. Luo, T. Zhou, and J. Zhang, Power-law behavior of transcriptional bursting regulated by enhancer-promoter communication, Genome Res., **34**, 106 (2024).

[91] L. Huang, Z. Yuan, P. Liu, and T. Zhou, Effects of promoter leakage on dynamics of gene expression, BMC Syst. Biol. **9**, 16 (2015).

[92] X. Yang, Y. Chen, T. Zhou, and J. Zhang, Exploring dissipative sources of non-Markovian biochemical reaction systems, Phys. Rev. E **103**, 052411 (2021).

[93] Y. Shin, Rich phase separation behavior of biomolecules, Mol. Cells **45**, 6 (2022).

[94] X. Kang and C. Li, A dimension reduction approach for energy landscape: Identifying intermediate states in metabolism-EMT network, Adv. Sci. **8**, 2003133 (2021).

[95] Y. Luo, G. Jiang, T. Yu, Y. Liu, L. Vo, H. Ding, Y. Su, W. W. Qian, H. Zhao, and J. Peng, ECNet is an evolutionary context-integrated deep learning framework for protein engineering, Nat. Commun. **12**, 5743 (2021).

[96] K. L. Saar, D. Qian, L. L. Good, A. S. Morgunov, R. Collepardo-Guevara, R. B. Best, and T. P. J. Knowles, Theoretical and data-driven approaches for biomolecular condensates, Chem. Rev. **123**, 8988 (2023).