

CONTROLLING STOCHASTICITY IN EPITHELIAL-MESENCHYMAL TRANSITION THROUGH MULTIPLE INTERMEDIATE CELLULAR STATES

CATHERINE HA TA, QING NIE AND TIAN HONG *

Department of Mathematics, Univ. of California Irvine
Irvine, CA 92697-3875, USA

ABSTRACT. Epithelial-mesenchymal transition (EMT) is an instance of cellular plasticity that plays critical roles in development, regeneration and cancer progression. Recent studies indicate that the transition between epithelial and mesenchymal states is a multi-step and reversible process in which several intermediate phenotypes might coexist. These intermediate states correspond to various forms of stem-like cells in the EMT system, but the function of the multi-step transition or the multiple stem cell phenotypes is unclear. Here, we use mathematical models to show that multiple intermediate phenotypes in the EMT system help to attenuate the overall fluctuations of the cell population in terms of phenotypic compositions, thereby stabilizing a heterogeneous cell population in the EMT spectrum. We found that the ability of the system to attenuate noise on the intermediate states depends on the number of intermediate states, indicating the stem-cell population is more stable when it has more sub-states. Our study reveals a novel advantage of multiple intermediate EMT phenotypes in terms of systems design, and it sheds light on the general design principle of heterogeneous stem cell population.

1. Introduction. Epithelial-mesenchymal transition (EMT) is an extreme form of cellular plasticity that is involved in morphogenesis, tissue regeneration and cancer progression. During EMT, epithelial cells undergo dramatic changes in cell morphology and behavior to form mesenchymal cells. These changes include loss of cell-to-cell junction, loss of cell polarity and acquisition of migratory and invasive properties [5] [16]. The migratory behavior of newly formed mesenchymal cells is critical for the formation of internal organs during embryonic development [5], and it is also involved in cancer metastasis, which often requires the dissemination of cancerous epithelial cells from the primary tumors [25]. After arriving to their destinations, mesenchymal cells sometimes revert to epithelial cells via a process called mesenchymal-epithelial transition (MET), suggesting that EMT is a reversible process [2].

Previous theories and experiments showed that EMT does not give rise to terminal mesenchymal phenotype in some biological scenarios [24] [30] [36]. In other words, intermediate epithelial-mesenchymal phenotype may be generated by partial EMT. Interestingly, it has been shown that partial EMT, or intermediate EMT phenotype, is associated with the stemness of epithelial and mesenchymal cells [9].

2010 *Mathematics Subject Classification.* 92C42, 93C05.

Key words and phrases. Noise attenuation, cellular plasticity, stem cell dynamics, heterogeneous cell population, multiple intermediate phenotypes, multi-step EMT.

* Corresponding authors: Qing Nie and Tian Hong.

In addition, it was hypothesized that partial EMT is responsible for collective migration and invasiveness of cancer cells [24]. These observations and hypothesis strongly suggest that EMT is a multi-step transition, and various degrees of the transition may have distinct physiological outcomes.

Using modeling and experimental approaches, we recently demonstrated the existence of multiple intermediate phenotypes in the EMT spectrum [12]. This is consistent with the observations which show four epigenetic states in multi-step EMT [29], multiple epithelial/mesenchymal stem cell phenotypes [1] and four distinct types of ovarian cancer cells in the EMT spectrum [14]. We also showed that the multiple intermediate states arise from complex gene regulatory networks consisting of interconnected feedback loops [12]. Previous theoretical studies suggested the role of heterogeneity in cell population and its feedback control in stem cell regeneration [20]. However, it is unclear why the epithelial/mesenchymal cell population needs these multiple intermediate states, or why the stem cells in the EMT spectrum need to have several subtypes. In particular, what performance objective does the system evolve to achieve by having multiple intermediate states at the expense of simplicity of cellular phenotypes?

To address these questions, we built and compared a series of mathematical models containing various number of intermediate states in EMT spectrum. By analyzing the behaviors of the system in facing noise in cell populations, we found that multiple intermediate phenotypes in the system help to attenuate the overall fluctuations of the cell population in terms of phenotypic compositions, suggesting their ability to stabilize a heterogeneous cell population. When we compared the models with at least one intermediate state, we found that the number of intermediate states positively correlates with the ability to attenuate noise on the intermediate states, indicating that the stem-cell population is more stable when it has more sub-states. These results suggest a performance objective that the EMT system might have evolved to achieve by having multiple intermediate phenotypes, and it sheds light on the general principles of heterogeneous stem cell population in term of systems design.

2. Mathematical models and stochastic simulation of multiple state EMT models.

2.1. Model construction. We explored the functional significance of the intermediate states in the EMT process by adapting the population dynamics to modelling the transition from epithelial to mesenchymal phenotypes. Each steady state in the transition was portrayed by a population of cells that assumed the phenotype reflective of that state. In other words, an n -state EMT system has n different populations of cell, with one of them being epithelial cells and another being mesenchymal cells. Hence there is a total of $n - 2$ groups that represent the cell populations of different intermediate stages. In this paper, we discussed the population dynamics for four different systems: two-state, three-state, four-state, and five-state EMT processes (Figure 1). In each system, each population of cells was characterized by death rate and transition rates from/to another population. The populations that corresponded to the intermediate states were given stem cell features, described by self-renewal rates and lower death-rates than other types of cells. We modeled the population of each cell phenotype with a linear ordinary differential equation (ODE), where we introduced a constant influx of cells to the mesenchymal population under the assumption that mesenchymal cells are mobile and invasive. With

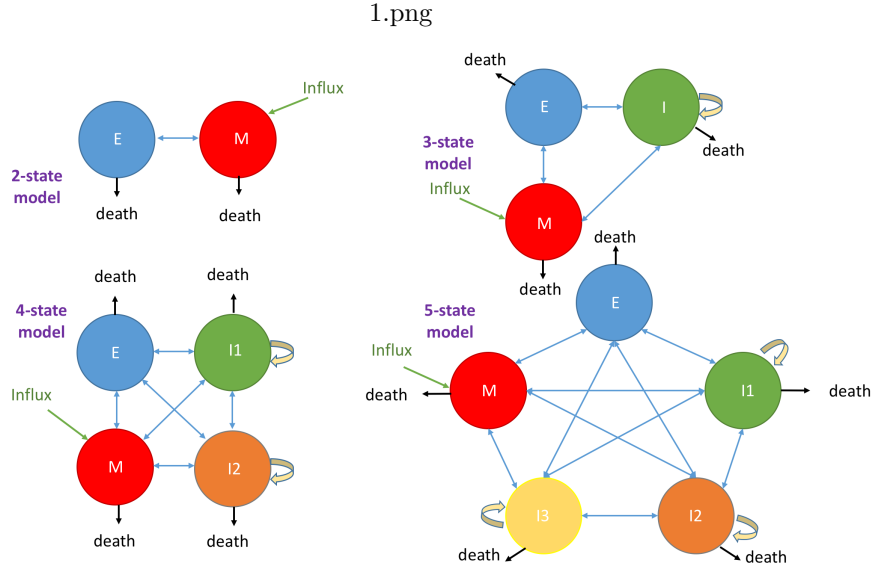


FIGURE 1. Four population dynamics models for four distinct EMT systems: two-state, three-state, four-state, and five-state transitions. In each model, E denotes the epithelial population, M mesenchymal population, and I_i the population corresponding to the i^{th} intermediate state. In each model, the blue arrows describe the transitions of cells between two populations, where the arrow heads describe the direction of each transition. The black arrows indicate cell death, the yellow arrows the self-renewal of (stem) cells in the intermediate states, and the green arrows that point towards the mesenchymal population indicate a constant influx of cells into this population.

this influx, all of the matrix systems in our models are nonsingular. The ODE that depicts the change in the population size of a particular cell phenotype P_i , $i = 1, 2, \dots, n$, is as followed:

$$\frac{dP_i}{dt} = \sum_{k \neq i} \alpha_{ki} P_k - \sum_{k \neq i} \alpha_{ik} P_i - d_i P_i + s_i P_i \tag{1}$$

where α_{ki} is the cellular conversion rate from P_k to P_i , d_i is the death rate, and s_i is the self-renewal rate where $s_i = 0$ if the phenotype of the cell population P_i is epithelial or mesenchymal. The cellular conversion rates, death rates, and self-renewal rates all assume constant values. In addition, if P_i describes the population of mesenchymal cell, then

$$\frac{dP_i}{dt} = \sum_{k \neq i} \alpha_{ki} P_k - \sum_{k \neq i} \alpha_{ik} P_i - d_i P_i + s_i P_i + I_M \tag{2}$$

with I_M being the constant influx of cells into the mesenchymal population. For our simulations, we chose the range for our parameter values to be between $1e-5$ and 10 , with the death rates d_i of the “stem cell” populations to be between $1e-5$ and $1e-2$ to reflect the lower death rates of stem cells. Meanwhile, since the epithelial

and mesenchymal populations were assumed to have higher death rates, we chose the death rates to be between 0.05 and 10. The quantity of cell population is in arbitrary unit. One time unit in our model corresponds to one day.

In order to explore the noise attenuation properties of each population dynamics model, we introduced multiplicative noise $\sigma_i P_i dW_i$ to our ODE system via multiple means, where σ_i is the noise coefficient and W_i describes the Standard Brownian motion. This noise is referred to as “input noise” in this paper. First, we added noise to the epithelial and mesenchymal populations, then observed the effects that noise introduction had on each individual population as well as overall. Next, we added noise to one intermediate population besides the two existing noises on epithelial and mesenchymal cells, and finally we introduced multiplicative noise to each cell population. To ensure consistency between the population fluctuations, we chose the noise coefficient $\sigma_i = 1$ for all $i = 1, 2, \dots, n$.

2.2. Model comparison through differential evolution algorithm. To compare various models with distinct number of parameters, we used Differential Evolution (DE) to optimize these models with respect to their ability to attenuate noise, and we subsequently compared the optimized parameter sets in terms of the noise attenuation property. DE is a method for improving a set of parameter values (i.e., a parameter vector \mathbf{x} , as above) with respect to an objective function, $\Phi(\mathbf{x})$, by generating a sequence of trial parameter vectors by processes of reproduction and selection [26]. Reproduction generates an offspring parameter vector \mathbf{u} from a parent parameter vector \mathbf{x} by diversification. If the offspring performs better than the parent, then the cost value produced by the objective function for the offspring is less than that of the parent, i.e. $\Phi(\mathbf{u}) < \Phi(\mathbf{x})$. In this case, the parent vector \mathbf{x} is replaced by the offspring vector \mathbf{u} in the next generation. The DE algorithm was previously shown to be efficient in optimizing models for biological systems [13]. With this algorithm, we iteratively searched for a newer set of parameter values that provided a better cost to the objective function $\Phi(x)$ than the previous set until the cost converged to the most optimal value. We chose the cost to be the average coefficient of variation of the population size for all phenotypes, i.e. the sum of the coefficients of variation of all the cellular populations divided by the number of states in a particular EMT system that we study. As a result, a lower cost value implies an overall reduction in fluctuations in each cell population and is thus desirable. We applied the optimization algorithm ten times to obtain ten sets of parameters with their corresponding cost values.

During DE, parameter vectors evolve from generation to generation. Each generation (indexed by $t = 0, 1, \dots$) consists of N parameter vectors \mathbf{x}_j , $j = 1, \dots, N$. Hence, the real number $x_{ij}(t)$ is the value of the i^{th} parameter in the j^{th} parent in the t^{th} generation. Let $\mathbf{u}_j(t)$ be the parameter vector for the single offspring of the j^{th} parent in the t^{th} generation. The components of this vector, $u_{ij}(t)$ for $i = 1, \dots, D$, are constructed in two steps (mutation and crossover). Then, given the two parameter vectors $\mathbf{x}_j(t)$ and $\mathbf{u}_j(t)$, a decision is made as to which one is propagated to generation $t + 1$. The following three operations propagate parameter vectors from one generation to the next:

1. Mutation: First, we create a “mutant” vector $\mathbf{v}_j(t)$ by perturbing a parental vector $\mathbf{x}_j(t)$. In our DE approach, we let the perturbation vector be the difference between the parameter vectors of two additional parents, j' and j'' , chosen at random from the t^{th} generation of parents. All three parents must

be different. The “mutant” vector is defined by:

$$\mathbf{v}_j(t) = \mathbf{x}_j(t) + F \cdot (\mathbf{x}'_j(t) - \mathbf{x}''_j(t)) \tag{3}$$

where F is a scalar, $0 < F < 1$, that determines how “aggressive” the mutation is.

2. Crossover: Next we allow for crossover between the parental parameter set $\mathbf{x}_j(t)$ and the mutant parameter set $\mathbf{v}_j(t)$. Component-wise, the offspring vector $\mathbf{u}_j(t)$ receives a parameter value from the mutant vector with probability C (the crossover probability) or from the parent vector with probability $1 - C$:

$$u_{i,j}(t) = \begin{cases} v_{i,j}(t), & \text{if } \text{rand}(0,1) \leq C \\ x_{i,j}(t), & \text{otherwise} \end{cases} \tag{4}$$

$$i = 1, 2, \dots, D \text{ and } j = 1, 2, \dots, N.$$

We choose $C = 0.5$ so that parental values and mutant values have equal chances to be in the offspring vector.

3. Selection: Depending on their relative performances, $\mathbf{x}_j(t)$ or $\mathbf{u}_j(t)$ passes on to the next generation. There are two possibilities here. For “greedy” selection, the offspring replaces its parent if it is superior:

$$\mathbf{x}_j(t+1) = \begin{cases} \mathbf{u}_j(t), & \text{if } \Phi(\mathbf{u}_j(t)) < \Phi(\mathbf{x}_j(t)) \\ \mathbf{x}_j(t), & \text{otherwise} \end{cases} \tag{5}$$

For “non-greedy” selection, the condition for replacement is:

$$\Phi(\mathbf{u}_j(t)) \leq \Phi(\mathbf{x}_j(t)) \tag{6}$$

We implement DE using the Matlab code available in [Price and Storn’s algorithm for DE \[38\]](#).

2.3. Application of fluctuation dissipation theorem. To measure the fluctuations on each cell population, we calculated the coefficient of variation of the population size for each cell phenotype from the covariance matrix that we found using the Fluctuation Dissipation Theorem [17] [18]. The Fluctuation Dissipation Theorem helps to predict the behaviors of a system in thermal equilibrium by establishing a relation between the frictional force and random force created by the Brownian particle’s impacts with surrounding molecules. We applied the version of the Fluctuation Dissipation Theorem mentioned in [32] to our system of linear stochastic ODEs:

$$d\mathbf{P} = A\mathbf{P}(t)dt + \Gamma(\mathbf{P}(t))d\mathbf{W}_t \tag{7}$$

where A is a constant matrix and $\Gamma(\mathbf{P}(t))d\mathbf{W}_t$ represents the multiplicative noise term. $\Gamma(\mathbf{P}(t))$ is a diagonal matrix whose i^{th} entry is $\sigma_i P_i(t)$. With this application of the Fluctuation Dissipation Theorem, we found the covariance matrix Σ for the equilibrium system through the formula

$$A\Sigma + \Sigma A^T = -\Gamma^2(\mathbf{P}_{ss}(t)) \tag{8}$$

with $\mathbf{P}_{ss}(t)$ being the steady states of the deterministic system $d\mathbf{P}/dt = A\mathbf{P}(t)$.

To quantify the noise attenuation performance of each EMT system, we aggregate the results of ten different optimization simulations using the DE algorithm [26] and calculate the mean and standard deviation of the coefficients of variation that describe the fluctuations in a specific population. Figures 2-8.

2.4. Sensitivity analysis. In our sensitivity analysis, we defined our parameters as followed: k_{ij} describes the cellular transition rate from the population in the i -state to the population in the j -state where $i, j = E, M, I, 1, 2,$ and 3 . Here, E denotes epithelial state, M mesenchymal state, I the intermediate state in the three-state EMT system, and 1, 2, and 3 denote the three intermediate states $I_1, I_2,$ and I_3 in the four-state and five-state systems. The parameter k_{sI} is the self-renewal rate of the intermediate population in the three-state EMT system while the parameters $k_{s1}, k_{s2},$ and k_{s3} are the self-renewal rates of the $I_1, I_2,$ and I_3 populations in the five-state process.

The sensitivity of each parameter is measured by the average change in the average coefficient of variation upon perturbation by 50%. For some transition rates, a positive change in the parameter value due to perturbation results in a significantly smaller average coefficient of variation but this parameter value can be discounted because it lies outside the designated value range that we used in our optimization algorithm Figures 4 and 8.

3. Results.

3.1. Intermediate states reduce stochastic fluctuations in EMT. We first asked if the intermediate state helps to attenuate the fluctuations in E and M cells. We used two optimized parameter sets for a two-state (E and M) model and a three-state (E, I and M) model for comparison, and we obtained two representative trajectories under the influence of identical level of input noise in E and M states. When we compared these two trajectories in terms of the fractions of M cells (Figure 2A), we found that the fluctuations in the fraction of M cells were reduced by 3.6% after adding the intermediate state. To exclude the possibility that the reduction of fluctuations is due to the lack of input noise in the intermediate state, we compared the two-state model and the three-state model with noise in all of the three states, and the latter model still showed discernible reduction of fluctuations in fraction of M cells under this condition (Figure 2B). We performed similar analysis on the fraction of E cells, and we found that when the input noise was only in E and M states, the three-state model performed significantly better than the two-state in terms of minimizing the fluctuations in E cells (Figure 3A), but when the input noise was in all the states, adding intermediate state resulted in a moderate increase of the fluctuations (Figure 3B). These results suggest that the intermediate state attenuates the fluctuations in M cells, and it may decrease or increase the fluctuations in E state depending on the source of noise.

We quantified the average fluctuations of all cellular states with the two-state and the three-state models. When we added noise on E and M states, the average cellular fluctuations were markedly reduced in the presence of the intermediate state (Figure 4A), and there was still a moderate reduction of average fluctuations when the noise was on all cellular states (Figure 4B). This suggests that the overall performance of the three-state model is better than that of the two-state model in terms of stability of phenotypic compositions.

We next asked which cellular state transition rates are critical for maintaining the stability. By perturbing the parameters representing the rates transitions among the three cellular states (input noise was added to all three states), we found that the performance of noise attenuation in the three-state system is most sensitive to three parameters: the transition rate from I to E (k_{IE}), the transition rate from I to

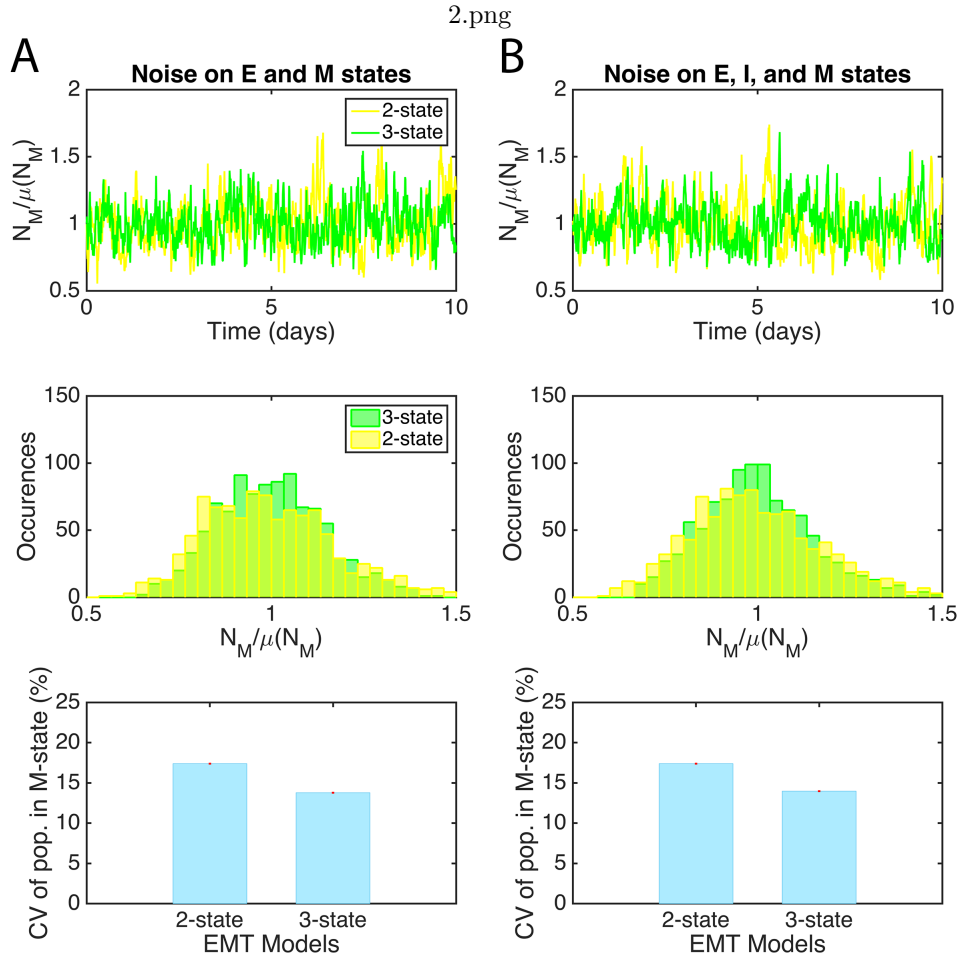


FIGURE 2. Comparison of the noise attenuation property between the two-state and three-state EMT systems in terms of the effects on the mesenchymal population size. **A)** Multiplicative noise is introduced to the epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, intermediate (three-state system), and mesenchymal populations. Top two panels are the time-course trajectories that represent the normalized number of mesenchymal cells $N_M/\mu(N_M)$ over a period of 10 days. To obtain the normalization, we perform stochastic simulations on the steady state population of mesenchymal cells, then divide the mesenchymal population size at each time point by its mean value obtained over the 10-day period. In the middle two panels, the normalized mesenchymal population size is plotted against the number of times that particular size occurs. Yellow: two-state EMT, green: three-state EMT. Bottom two panels display the quantification of the noise attenuation performance of the two-state and three-state systems using the mean and standard deviation of the coefficients of variation (CV) of the mesenchymal population. The mean is plotted here in the form of a bar graph (blue), while the standard deviation is described by the red error bar.

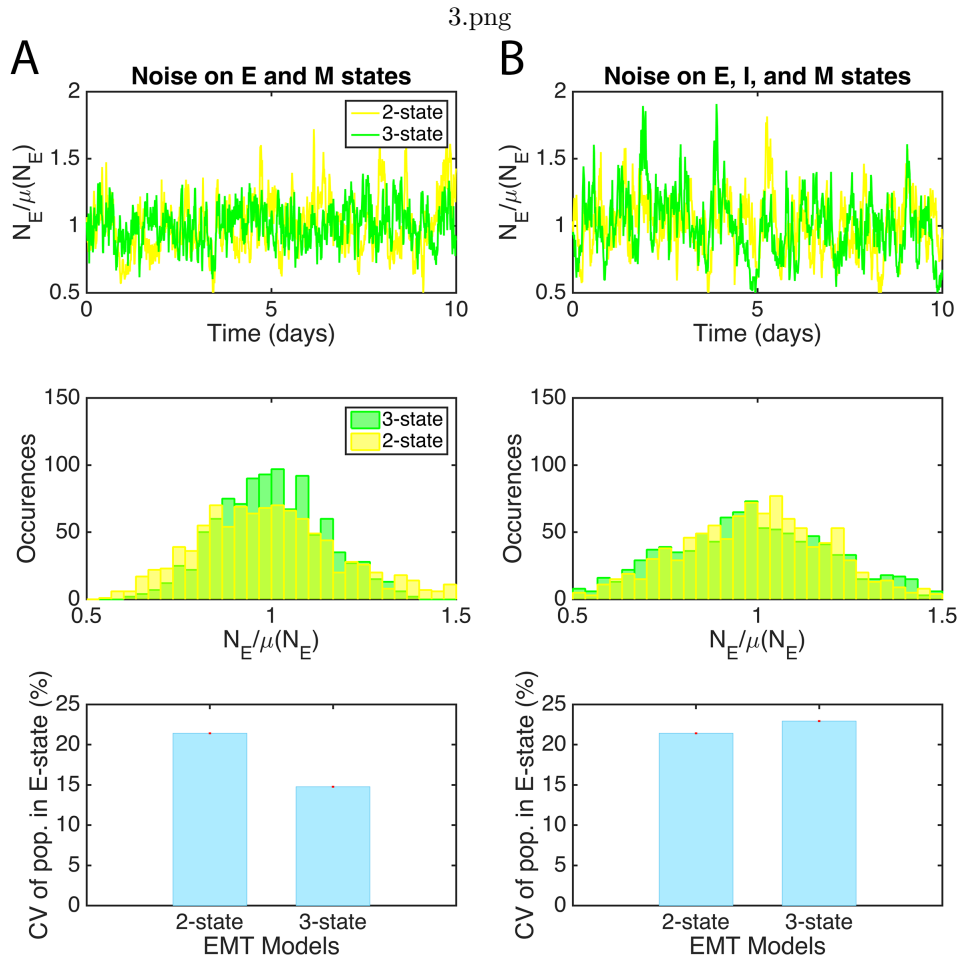


FIGURE 3. Comparison of the noise attenuation property between the two-state and three-state EMT systems in terms of the effects on the epithelial population size. **A)** Multiplicative noise is introduced to epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, intermediate (three-state system), and mesenchymal populations. Top two panels are the time-course trajectories that represent the normalized number of epithelial cells $N_E/\mu(N_E)$ over a period of 10 days. Middle two panels illustrate the distribution of the different population sizes of the normalized epithelial population. Here, the normalized epithelial population size is plotted against the number of times that particular size occurs. The color coding scheme for each system is similar to that of Figure 2. Bottom two panels display the quantification of the noise attenuation performance of the two-state and three-state systems.

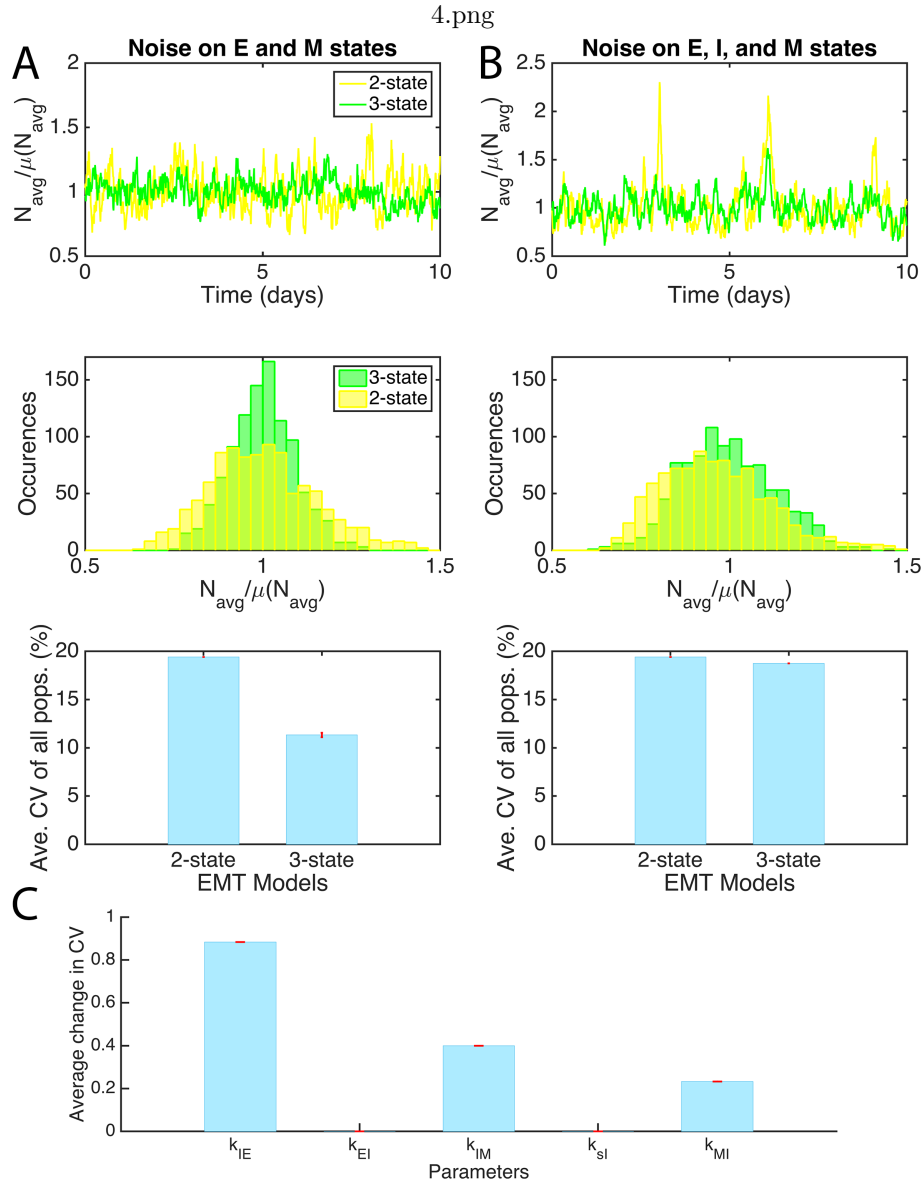


FIGURE 4. Comparison of the noise attenuation property between the two-state and three-state EMT systems on the average population size. **A)** Noise is introduced to epithelial and mesenchymal populations. **B)** Noise is introduced to the epithelial, intermediate (three-state system), and mesenchymal populations. Top panels: trajectories of normalized average number of cells $N_{avg}/\mu(N_{avg})$ over 10 days. Middle panels: distribution of the different sizes of the normalized average population using the color coding scheme of Figure 2. Bottom panels: quantification of the noise attenuation performance of both systems. **C)** Sensitivity analysis of the parameters representing unique cell transition rates in the three-state EMT system. Here, we plot the mean of the average change in the average CV as bar graphs (blue) accompanied by red error bars that describe the standard deviation of the average change.

M (k_{JM}) and the transition rate from M to I (k_{MI}) (Figure 4C). These results suggest that the transitions involving the intermediate states are important for noise attenuation, and they further corroborate the advantage of the intermediate state. They correlation between these transition rates and the ability of the system to attenuate noise may be tested in future experiments.

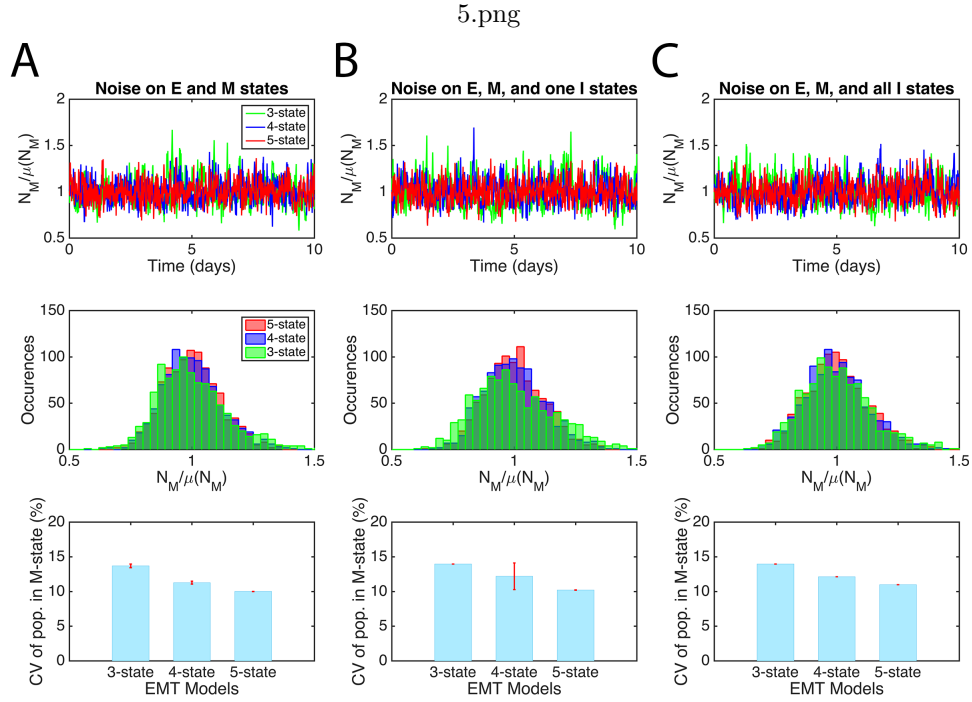


FIGURE 5. Comparison of the noise attenuation property between the three-state, four-state, and five-state EMT systems in terms of the effects on the mesenchymal population size. **A)** Multiplicative noise is introduced to epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, mesenchymal, and one intermediate populations. **C)** Multiplicative noise is introduced to all the populations. Top three panels are the time-course trajectories that represent the normalized number of mesenchymal cells $N_M/\mu(N_M)$ over a period of 10 days. To obtain the normalization, we perform similar stochastic simulations to those in Figures 2-4. Middle three panels illustrate the distribution of the different population sizes of the normalized mesenchymal population. Here, the normalized mesenchymal population size is plotted against the number of times that particular size occurs. Green: three-state EMT, blue: four-state EMT, red: five-state EMT. Bottom three panels display the quantification of the noise attenuation performance of the three-state, four-state, and five-state systems.

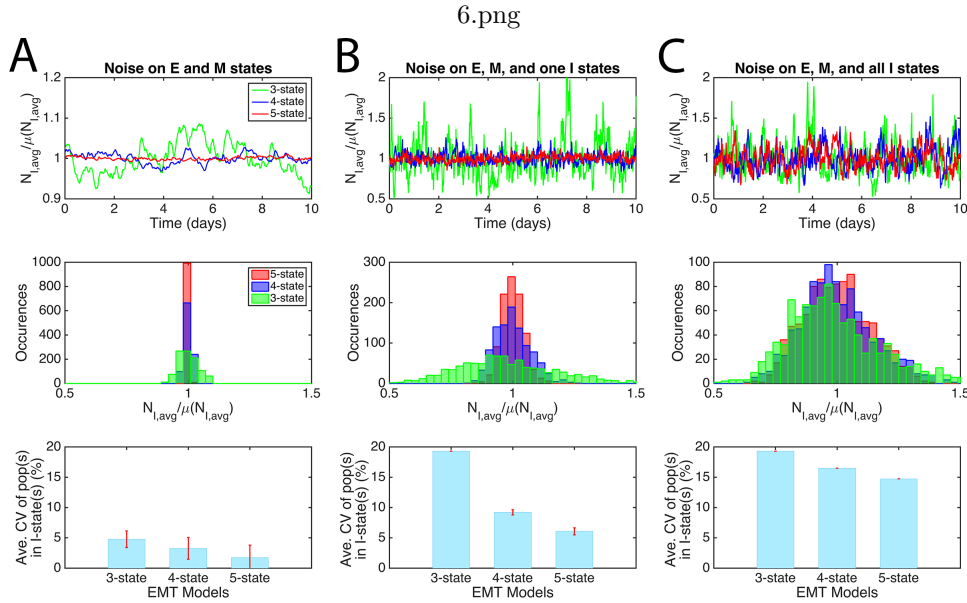


FIGURE 6. Comparison of the noise attenuation property between the three-state, four-state, and five-state EMT systems in terms of the effects on the average intermediate population size. **A)** Multiplicative noise is introduced to epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, mesenchymal, and one intermediate populations. **C)** Multiplicative noise is introduced to all the populations. Top three panels are the time-course trajectories that represent the normalized average number of cells taken over all the intermediate states $N_{I,avg}/\mu(N_{I,avg})$ over a period of 10 days. Middle three panels illustrate the distribution of the different population sizes of the normalized average intermediate population. Here, the normalized average intermediate population size is plotted against the number of times that particular size occurs. The color coding scheme for each system is similar to that of Figure 5. Bottom three panels display the quantification of the noise attenuation performance of the three-state, four-state, and five-state systems.

3.2. Multiple intermediate states enhance noise attenuation. Next, we explored the advantages of modeling the EMT system with multiple steady states. Using identical input noise on E and M populations, we examined the noise buffering property of three different EMT models: three-state, four-state, and five-state transitions. Analysis of the fluctuations in M cells demonstrates a reduction in noise effects in systems with more than one intermediate states, as evident by the normalized trajectories of the fraction of M cells and the coefficient of variation that reflects the variability of the mesenchymal population size (Figure 5A). In addition, five-state system filters noise more efficiently than four-state system. Similarly, when we introduced noise to one intermediate population in addition to the E and M

7.png

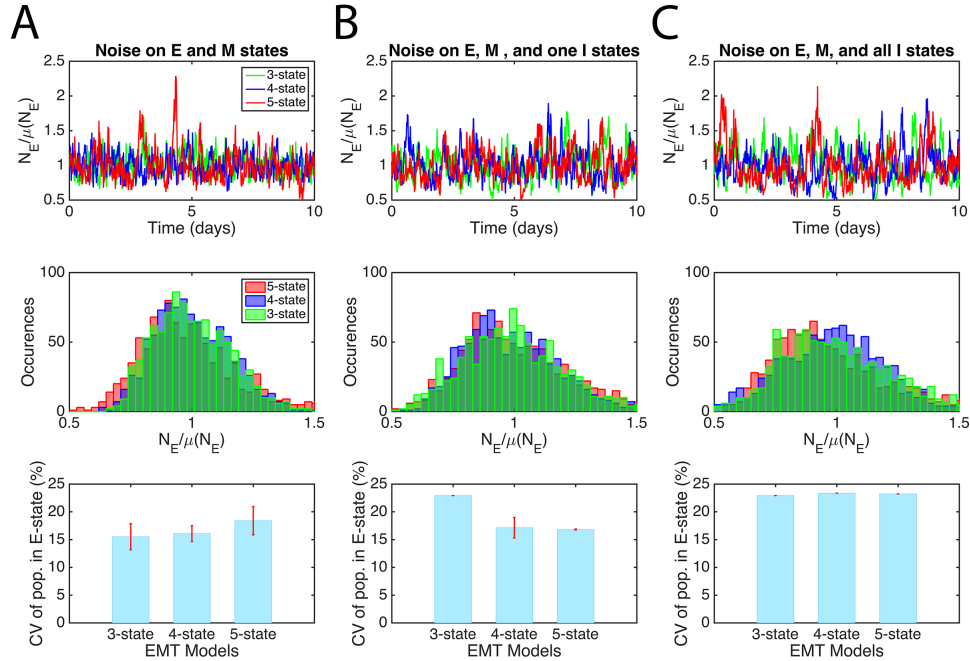


FIGURE 7. Comparison of the noise attenuation property between the three-state, four-state, and five-state EMT systems in terms of the effects on the epithelial population size. **A)** Multiplicative noise is introduced to epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, mesenchymal, and one intermediate populations. **C)** Multiplicative noise is introduced to all the populations. Top three panels are the time-course trajectories that represent the normalized number of epithelial cells $N_E/\mu(N_E)$ over a period of 10 days. Middle three panels illustrate the distribution of the different population sizes of the normalized epithelial population. Here, the normalized epithelial population size is plotted against the number of times that particular size occurs. The color coding scheme for each system is similar to that of Figure 5. Bottom three panels display the quantification of the noise attenuation performance of the three-state, four-state, and five-state systems.

populations, we observed the same correlation between the number of intermediate populations in an EMT system and the noise attenuation property of that system (Figure 5B). We also confirmed this result by adding noise to all the populations in every EMT system, where the five-state system achieved the best noise filtering results, followed closely by the four-state system (Figure 5C). From analyzing the responses of the average intermediate population to noise with similar approach, we found that the populations of the five-state and four-state systems outperformed

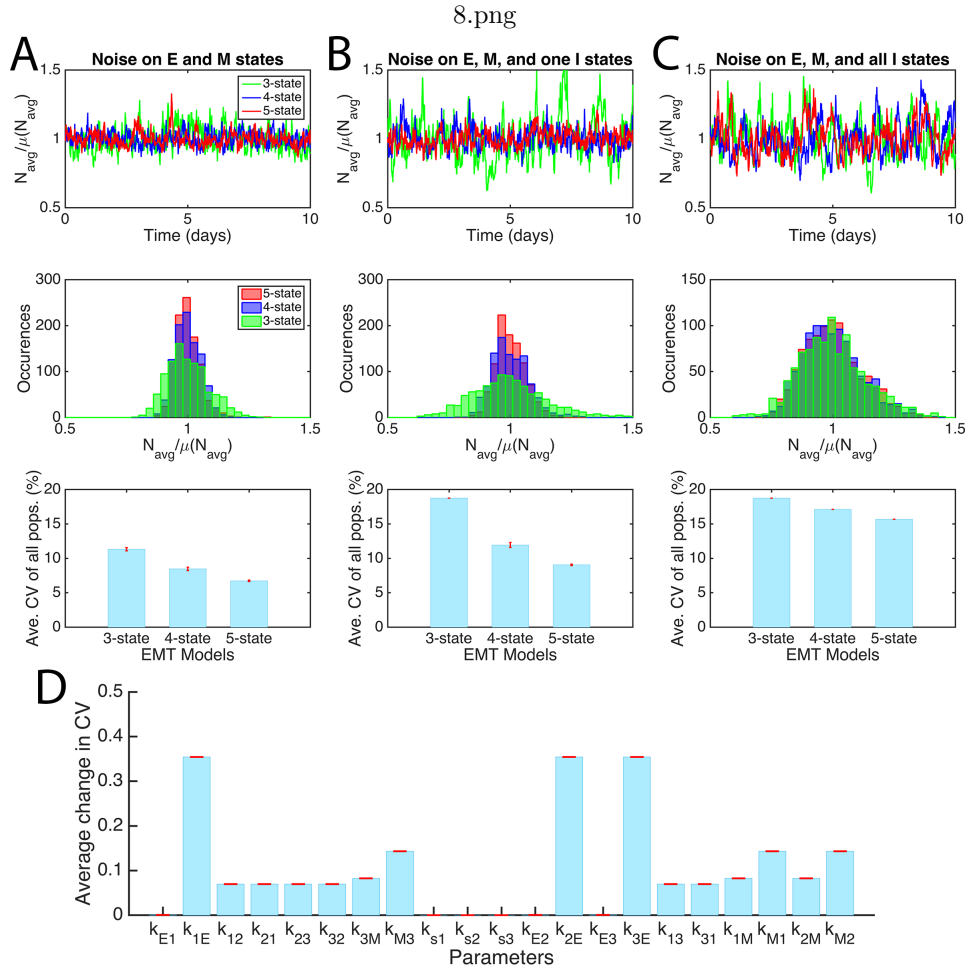


FIGURE 8. Comparison of the noise attenuation property between the three-state, four-state, and five-state EMT systems in terms of the effects on the average population size. **A)** Multiplicative noise is introduced to epithelial and mesenchymal populations. **B)** Multiplicative noise is introduced to the epithelial, mesenchymal, and one intermediate populations. **C)** Multiplicative noise is introduced to all the populations. Top three panels are the time-course trajectories that represent the normalized average number of cells $N_{avg}/\mu(N_{avg})$ over a period of 10 days. Middle three panels illustrate the distribution of the different population sizes of the normalized average population. The color coding scheme for each system is similar to that of Figure 5. Bottom three panels display the quantification of the noise attenuation performance of all the EMT systems. **D)** Sensitivity analysis of the parameters representing unique cell transition rates in the five-state EMT system. Here, we plot the mean of the average change in the average CV as bar graphs (blue) accompanied by red error bars that describe the standard deviation of the average change.

that of the three-state system in attenuating noise at the intermediate levels, regardless of how noise was introduced to the different populations (Figure 6A-C). This suggests that the number of stem cell sub-states correlates with their ability to reduce the fluctuations of stem cell population. When we studied the effects of noise on the fraction of E cells, we did not observe a similar trend to that of the M and I cells. When input noise was added to E and M populations only, the epithelial population of the five-state system was more susceptible to the fluctuations (Figure 7A) while the epithelial population of the same system showed the least variability when noise was also added to one intermediate population (Figure 7B). With noise in every state, the difference in the performance of each system in filtering noise at the epithelial level was not discernible (Figure 7C). These results suggest that when the number of intermediate state varies, there might be a tradeoff between attenuating noise in E cells and other cells.

Lastly, we quantified the overall performance of each EMT system by considering the fluctuations on the average population of all cellular states. In the scenario where noise is only introduced to the E and M states, the five-state system performs the best at suppressing noise, while the four-state system displays better noise attenuation than the three-state system (Figure 8A). With the addition of noise in one intermediate state, the four-state and the five-state EMT processes demonstrate improvement in noise attenuation from their three-state counterpart by 6.9% and 9.6% respectively (Figure 8B). Likewise, we confirmed our results with the addition of noise to all cellular populations and concluded that the EMT system achieves better noise attenuation property when more intermediate states are taken into consideration (Figure 8C).

Using the five-state EMT process with noise in every state, we analyzed the sensitivity of the parameters that described the cellular transition/self-renewal rates. For our analysis, we examined the average change in the average coefficient of variation of all five populations upon perturbing each parameter. We performed the optimization procedure ten times, whereupon, we carried out the perturbations each time and presented our results in Figure 8D. We found that the parameters that described the transition rate from any intermediate population I_1 , I_2 , or I_3 to the epithelial population (k_{1E} , k_{2E} , k_{3E}) are most sensitive to their own perturbations, resulting in a marked increase in the average coefficient of variation, therefore a decrease in noise buffering ability of this EMT system. Besides those three transition rates, the cellular transitions from the mesenchymal population to the intermediate populations (k_{M1} , k_{M2} , k_{M3}), also show considerable sensitivity to perturbations. We thus recognize the crucial roles that the intermediate states such as the I_1 , I_2 , and I_3 states play in ensuring the better noise attenuation performance of the five-state system. The significance of the transitions between intermediate states and the others reinforce the notion that having multiple intermediate states is beneficial to reducing the effects of the fluctuations on the overall population.

4. Discussion. Studying the multi-step and reversible dynamics of EMT is essential for understanding the roles of EMT in various biological processes such as development and cancer progression. Previous studies revealed multiple intermediate states lying between terminal epithelial and mesenchymal states [12] [14] [29], and this is consistent with the observations that epithelial cell populations show remarkable heterogeneity in normal and tumor tissues [19][35]. These intermediate states are associated with stemness and invasiveness during cancer progression

[9] [15], but the functions of these multiple intermediate states and the complex transitions among all the phenotypic states are unclear. In this study, we used mathematical models to show that systems with multiple intermediate (stem cell) states have advantages in attenuating fluctuations of the heterogeneous cell population, and this effect is particularly beneficial for maintaining stable fractions of stem cells in cell populations, thus providing a strategy for maintaining homeostasis in facing stochasticity of cell fate changes. The possible performance objective at systems level suggests a new design principle for multi-step EMT and other multi-state systems involving transitions among diverse types of cells.

Complex networks in biological systems have been extensively studied in terms of design principles and performance objectives [21] [31]. However, most of these studies focused on intracellular networks of molecular interactions and influence, and much less is known for networks formed by phenotypic transitions among cell types [10]. We showed that the later type of networks can carry rich information in terms of performance objectives such as facilitating noise attenuation. This finding improves our understanding of systems design at tissue level, and suggests that the formation of cellular states and the transitions among them may have evolved to reach performance objectives that are beneficial to the organisms. It is likely that multiple performance objectives are influencing the evolution of the cellular networks, and further studies are needed to identify these objectives.

Due to the stochastic nature of biochemical reactions and fluctuating environmental conditions, many biological systems are designed to attenuate noise such that stable response to signals can be attained [6] [33]. Intermediate states and regulators have been previously found to improve the robustness of developmental patterning [3]. Our finding that adding intermediate cell states can reduce the fluctuations in cell population provides another strategy for noise attenuation at tissue level. The fluctuations that we included in our simulations represent the stochasticity of cell division [7], cell death [8] and phenotypic transitions [4], all of which were previously observed in experiments. Nonetheless, the sources of these types of noise are mostly molecular fluctuations that affect various types of rates of the cellular activities [37]. In order to describe the dynamics of the noise in more details, future models will need to incorporate the stochastic dynamics of biomolecules into the framework.

Multiple intermediate cell states may arise from complex gene regulatory networks with interconnected positive feedback loops [12] [13]. It is conceivable that the formation of more intermediate cells would require more complex gene regulatory networks, which in turn need some other strategies and/or more energy to control. Therefore, the correlation between the number of states and the ability to attenuate noise suggests that a tradeoff between the simplicity of the gene regulatory network and noise attenuation may exist when the system is subject to design via evolution.

Dynamic equilibrium of sub-states of stem cells is observed not only in the EMT system, but also in hematopoietic and embryonic stem cell populations [11] [27] [34]. Moreover, it was suggested that heterogeneity of cell population might be related to disseminated cancer cell dormancy [28]. Although these sub-states of stem cells might represent distinct functional entities in specific contexts, the recurring phenomena indicate that there might be advantages of this dynamic behavior at systems level. Our modeling work provides a general framework to study the cell differentiation systems with multiple sub-states of stem cells, and our conclusions

with respect to the noise attenuation property of multiple stem cell states can apply to other systems, such as embryonic stem cell and disseminated cancer cell [11] [28]. In addition, feedback control among multiple cellular states have been shown to be critical to the growth and stability of cell populations [20] [22] [23]. Future works are warranted to investigate the role of feedback controls involving the multiple intermediate states.

In conclusion, our computational analysis shows that the existence of multiple intermediate states between epithelial and mesenchymal states and the transitions among these states are able to attenuate the fluctuations of the fractions of cell population. We found a general correlation between the number of states and the ability to attenuate fluctuations. In particular, the fluctuations of the stem cell populations are reduced by increasing the number of intermediate states. These results improve the understanding of the intermediate states in the EMT system in terms of performance objectives, and provide insights into the stem cell systems with multiple sub-states in general.

Acknowledgments. The authors would like to thank Dr. Xing Dai for stimulating discussions. The work is partially supported by National Institute of Health grants R01GM107264, R01NS095355, P50GM76516, and R56AR064532, and the National Science Foundation grants DMS1161621 and DMS1562176.

REFERENCES

- [1] J. Baulida and A. Garcia de Herreros, [Snail1-driven plasticity of epithelial and mesenchymal cells sustains cancer malignancy](#), *Biochim Biophys Acta*, **1856** (2015), 55–61.
- [2] T. Brabletz, A. Jung and S. Reu, et al. [Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment](#), *Proc Natl Acad Sci U S A*, **98** (2001), 10356–10361.
- [3] A. Q. Cai, K. Radtke and A. Linville, et al. [Cellular retinoic acid-binding proteins are essential for hindbrain patterning and signal robustness in zebrafish](#), *Development*, **139** (2012), 2150–2155.
- [4] H. H. Chang, M. Hemberg and M. Barahona, et al. [Transcriptome-wide noise controls lineage choice in mammalian progenitor cells](#), *Nature*, **453** (2008), 544–547.
- [5] J. Chen, Q. Han and D. Pei, [EMT and MET as paradigms for cell fate switching](#), *J Mol Cell Biol*, **4** (2012), 66–69.
- [6] M. Chen, L. Wang and C. C. Liu, et al. [Noise attenuation in the ON and OFF states of biological switches](#), *ACS Synth Biol*, **2** (2013), 587–593.
- [7] S. Di Talia, J. M. Skotheim and J. M. Bean, et al. [The effects of molecular noise and size control on variability in the budding yeast cell cycle](#), *Nature*, **448** (2007), 947–951.
- [8] S. Gaudet, S. L. Spencer and W. W. Chen, et al. [Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis](#), *PLoS Comput Biol*, **8** (2012), e1002482.
- [9] A. Grosse-Wilde, A. Fouquier d’Herouel and E. McIntosh, et al. [Stemness of the hybrid epithelial/mesenchymal state in breast cancer and its association with poor survival](#), *PLoS One*, **10** (2015), e0126522.
- [10] Y. Hart, Y. E. Antebi and A. E. Mayo, et al. [Design principles of cell circuits with paradoxical components](#), *Proc Natl Acad Sci U S A*, **109** (2012), 8346–8351.
- [11] K. Hayashi, S. de Sousa Lopes and F. Tang, et al. [Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states](#), *Cell Stem Cell*, **3** (2008), 391–401.
- [12] T. Hong, K. Watanabe and C. Ta, et al. [An *ovol2-zeb1* mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states](#), *PLoS Comput Biol*, **11** (2015), e1004569.
- [13] T. Hong, C. Oguz and J. J. Tyson, [A mathematical framework for understanding four-dimensional heterogeneous differentiation of CD4+ T cells](#), *Bulletin of Mathematical Biology*, **77** (2015), 1046–1064.

- [14] R. Y. Huang, M. K. Wong and T. Z. Tan, et al. [An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib \(AZD0530\)](#), *Cell Death Dis*, **4** (2013), e915.
- [15] M. K. Jolly, D. Jia and M. Boareto, et al. [Coupling the modules of EMT and stemness: A tunable 'stemness window' model](#), *Oncotarget*, **6** (2015), 25161–25174.
- [16] R. Kalluri and R. A. Weinberg, The basics of epithelial-mesenchymal transition, *The Journal of Clinical Investigation*, **119** (2009), 1420–1428.
- [17] J. Keizer, *Statistical Thermodynamics of Nonequilibrium Processes*, Springer-Verlag, 1987.
- [18] R. Kubo, [The fluctuation-dissipation theorem](#), *Reports on Progress in Physics*, **29** (1966), 255–284.
- [19] D. A. Lawson, N. R. Bhakta and K. Kessenbrock, et al. [Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells](#), *Nature*, **526** (2015), 131–135.
- [20] J. Lei, S. A. Levin and Q. Nie, [Mathematical model of adult stem cell regeneration with cross-talk between genetic and epigenetic regulation](#), *Proc Natl Acad Sci U S A*, **111** (2014), E880–E887.
- [21] W. A. Lim, C. M. Lee and C. Tang, [Design principles of regulatory networks: Searching for the molecular algorithms of the cell](#), *Mol Cell*, **49** (2013), 202–212.
- [22] X. Liu, S. Johnson and S. Liu, et al. *Nonlinear Growth Kinetics of Breast Cancer Stem Cells: Implications for Cancer Stem Cell Targeted Therapy*, Sci Rep, 2013.
- [23] W. C. Lo, C. S. Chou and K. K. Gokoffski, et al. [Feedback regulation in multistage cell lineages](#), *Math Biosci Eng*, **6** (2009), 59–82.
- [24] M. Lu, M. K. Jolly and H. Levine, et al. [MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination](#), *Proc Natl Acad Sci U S A*, **110** (2013), 18144–18149.
- [25] S. A. Mani, W. Guo and M. J. Liao, et al. [The epithelial-mesenchymal transition generates cells with properties of stem cells](#), *Cell*, **133** (2008), 704–715.
- [26] K. V. Price, R. M. Storn and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Springer, Berlin, 2005.
- [27] Y. Shen, C. Shi and W. Wei, et al. [The heterogeneity and dynamic equilibrium of rat embryonic stem cells](#), *Cell Res*, **21** (2011), 1143–1147.
- [28] M. S. Sosa, P. Bragado and J. A. Aguirre-Ghiso, [Mechanisms of disseminated cancer cell dormancy: An awakening field](#), *Nat Rev Cancer*, **14** (2014), 611–622.
- [29] W. L. Tam and R. A. Weinberg, [The epigenetics of epithelial-mesenchymal plasticity in cancer](#), *Nature Medicine*, **19** (2013), 1438–1449.
- [30] X. J. Tian, H. Zhang and J. Xing, [Coupled reversible and irreversible bistable switches underlying TGF- \$\beta\$ -induced epithelial to mesenchymal transition](#), *Biophysical journal*, **105** (2013), 1079–1089.
- [31] J. J. Tyson and B. Novak, [Functional motifs in biochemical reaction networks](#), *Annu Rev Phys Chem*, **61** (2010), 219–240.
- [32] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd edition, Elsevier, Amsterdam, 2007.
- [33] L. Wang, J. Xin and Q. Nie, [A critical quantity for noise attenuation in feedback systems](#), *PLoS Comput Biol*, **6** (2010), e1000764, 17 pp.
- [34] W. Weston, J. Zayas and R. Perez, et al. [Dynamic equilibrium of heterogeneous and interconvertible multipotent hematopoietic cell subsets](#), *Sci Rep*, **4** (2014), 5199.
- [35] X. Ye, W. L. Tam and T. Shibue, et al. [Distinct EMT programs control normal mammary stem cells and tumour-initiating cells](#), *Nature*, **525** (2015), 256–260.
- [36] J. Zhang, X. J. Tian and H. Zhang, et al. *TGF- β -induced Epithelial-To-Mesenchymal Transition Proceeds Through Stepwise Activation of Multiple Feedback Loops*, Science Signaling, 2014.
- [37] L. Zheng, M. Chen and Q. Nie, [External noise control in inherently stochastic biological systems](#), *J Math Phys*, **53** (2012), 115616, 13 pp.
- [38] Differential Evolution (DE) for Continuous Function Optimization (an algorithm by Kenneth Price and Rainer Storn), Accessed in May 2016. Available from: <http://www1.icssi.berkeley.edu/~storn/code.html>.

Received February 2016; revised May 2016.

E-mail address: hmta@math.uci.edu

E-mail address: qnie@math.uci.edu

E-mail address: hong.tian@uci.edu