Exact solutions of the simplified March model for organizational learning

Hang-Hyun Jo *
Department of Physics, The Catholic University of Korea, Bucheon 14662, Republic of Korea

(Received 13 January 2024; accepted 2 July 2024; published 18 July 2024)

March's celebrated agent-based simulation model for organizational learning [J. G. March, Org. Sci. **2**, 71 (1991)] has been extensively studied in recent decades. Yet the model has not been fully understood due to the lack of analytical solutions of the model. We simplify the March model to take an analytical approach using master equations. We then derive exact solutions for some of the simplest yet nontrivial cases, and perform numerical estimation of master equations for more complicated cases. Both analytical and numerical results are in good agreement with agent-based simulations. These results are also compared to those of the original March model. Our approach enables us to rigorously understand the results of the simplified model as well as the original model, to a large extent.

DOI: 10.1103/PhysRevE.110.014303

I. INTRODUCTION

March introduced an agent-based simulation model for organizational learning in his seminal paper in 1991 [1]. Since then, the original March model, its variants, and other similar models have been extensively studied in recent decades [2–15]. March's model considers an external reality, an organizational code (code hereafter), and individual members of the organization. The code represents a set of norms, rules, etc. that is updated using the knowledge of superior individuals about the reality. Here superior individuals are those who have more correct knowledge about the reality than the code. On the other hand, all individuals also learn from the code about the reality. By doing so, the organizational knowledge about the reality is collected from individuals and is disseminated to them at the same time. March studied the effects of learning rates of individuals and of the code on their achieved knowledge about the reality. In order to consider more realistic situations, he took personnel turnover and environmental turbulence into account in his model to find that there may exist an optimal turnover rate that maximizes the achieved knowledge, depending on the learning rates.

We remark that the March model can be considered in the framework of opinion dynamics in networks [16–24]. That is, the code plays a hub node in a hub-and-spoke network, while individuals are dangling nodes [25]. Nodes update their opinions or beliefs according to their neighbors' opinions or beliefs, while the external reality plays an external field or source affecting all nodes. Asymmetric positions of the code and individuals in the network as well as partly directed influence between them make the model comparable to some opinion dynamics models considering the hierarchical structure of individuals [26–28]. In this sense, the March model can be seen as a hierarchical opinion dynamics model defined in a hub-and-spoke network. Overall, it implies that various analytical approaches developed for the opinion dynamics can be applied to the March model.

The March model and its variants have provided insights into management and business administration [8], but mostly by means of computer simulations of models [14]. It is probably because the computer simulation results are already enough to draw meaningful conclusions in the mentioned fields. However, in general, for a rigorous understanding of models, the derivation of their exact, analytical solutions is of the utmost importance. Analyzing the models at the most fundamental level helps us to precisely understand the mathematical structure of the dynamics. In our work, we simplify March's original model to explicitly write master equations describing the dynamics of the model. Then we derive exact solutions of the simplified model for some of the simplest, yet nontrivial cases. Numerical estimation of master equations is performed for the more complicated cases. Both analytical and numerical results are shown to be in good agreement with agent-based simulation results. Thus, our approach enables us to rigorously understand the results of not only the simplified model, but also the original model, to a large extent.

The paper is organized as follows. In Sec. II, we describe the original March model and our simplified version. In Sec. III, analytical, numerical, and simulation results of the simplified model are presented and compared to the results of the original March model. Finally, we conclude our work in Sec. IV.

II. MODELS

A. Original March model

As mentioned, the original model by March considers the external reality, the code, and individual members of the organization [1]. We remark that in this subsection, we will use

*Contact author: h2jo@catholic.ac.kr

mathematical symbols originally used in the March paper, and they are not supposed to be confused with symbols in the next subsection and throughout the paper. The model is based on the following assumptions:

(i) The reality is characterized in terms of an m-dimensional vector, each element of which may have the value of 1 or -1 with equal probabilities.

(ii) The code and n individuals in the organization have beliefs about the reality. The belief is also represented by an m-dimensional vector, each element of which may have the value of 1, 0, or -1 with equal probabilities. These beliefs may change over time.

(iii) At each time step, each individual may change elements of its belief that are different from those of the code unless the code's element is 0. Each of such elements of the individual changes to that of the code with a probability p_1, independently of other elements.

(iv) At the same time, the code updates its belief based on the beliefs of some individuals. For this, individuals whose beliefs are closer to the reality than the code is to the reality are identified, which is called a superior group. Then each element of the code's belief changes to the dominant element within the superior group with a probability p_2, independently of other elements.

So far, the reality has been assumed to be fixed and the individuals are not replaced by new ones. Thus, it is called a closed system. March first considered a homogeneous population in the closed system, in which all individuals are assigned the identical learning probability. Then he considered the heterogeneous population in the closed system, such that some individuals have higher learning probability than the others. Finally, a homogeneous population in the open system is also considered; in the open system, individuals may be replaced by new ones (turnover) and/or the reality changes over time (turbulence). The turnover probability is denoted by p_3 and the turbulence probability is by p_4. That is, with a probability p_3, each individual is replaced by a new one having a random belief vector at each time step. Also, each element of the reality shifts to the other value, i.e., from 1 to -1 or from -1 to 1, with a probability p_4.

It should be noted that the description of the original March model may not lead to a unique implementation for agent-based simulations [14]. For example, it is not clear within a time step whether the code learns from the superior individuals before the individuals learn from the code, or after the individuals learn from the code.

B. Simplified March model

Let us simplify the March model. As in the original model, we consider an external reality, a code, and N agents, as depicted in Fig. 1. At a time step t, the external reality, denoted by a variable $r(t)$, can have a value of 0 or 1. Beliefs of the code and agents about reality are respectively represented by variables, namely, $c(t) \in \{0, 1\}$ for the code and $\sigma_i(t) \in \{0, 1\}$ for the ith agent with $i = 1, \dots, N$. For a given initial condition of $r(0)$, $c(0)$, and $\{\sigma_i(0)\}$, each time step consists of four stages as follows:

(i) Every agent of $i \in \{1, \dots, N\}$ independently updates their belief by learning from the code with a *socialization* probability $p^{(i)} \in [0, 1]$:

$$\sigma_i(t+1) = c(t). \quad (1)$$

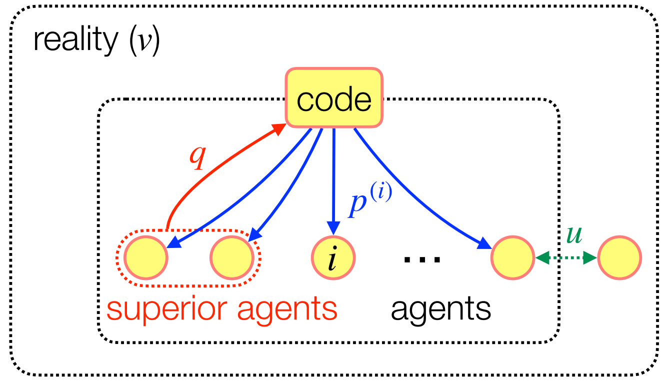

FIG. 1. Schematic diagram of the simplified March model that is composed of the reality, the code, and N agents. Blue arrows with probabilities $p^{(i)}$ for $i = 1, \dots, N$ denote the socialization of agents by learning from the code [Eq. (1)]. The red arrow with probability q denotes the codification of the code by learning from superior agents [Eq. (2)]. Agents can be replaced by new ones with probability u (green arrow), while the reality can change over time with probability v.

(ii) Each agent is replaced by a new agent with a *turnover* probability $u \in [0, 1]$, and the new agent is assigned a belief randomly drawn from $\{0, 1\}$.

(iii) The code learns from agents who are superior to the code. Here, superior agents indicate those whose beliefs are closer to the reality than the code's belief is. For example, if $r(t) = 1$, the code learns from superior agents only when $c(t) = 0$ and there is at least one agent with $\sigma_i(t) = 1$. Denoting a superior agent to the code by j, the code updates its belief with a *codification* probability $q \in [0, 1]$:

$$c(t+1) = \sigma_j(t) \text{ for } j \in \{i|\delta_{\sigma_i(t),r(t)} > \delta_{c(t),r(t)}\}, \quad (2)$$

where $\delta_{\cdot,\cdot}$ is a Kronecker delta.

(iv) With a *turbulence* probability $v \in [0, 1]$, the reality is assigned a new value randomly drawn from $\{0, 1\}$, which closes the time step.

Since the reality, the code, and the agents update their value or beliefs synchronously, the order of the four stages does not affect the result, except for the socialization and turnover of agents. Note that parameters $\{p^{(i)}\}$, q, u, and v in our simplified model correspond to p_1, p_2, p_3, and p_4 in the original March model [1], respectively.

Our simplified model can be called a closed system if $u = v = 0$; otherwise it is an open system. In the open system, the reality can change over time (turbulence) and/or agents can be replaced by new agents (turnover). In contrast, in the closed system, the reality is assumed to have the fixed value of 1 for the entire period of time, i.e., $r(t) = 1$ for all t, without loss of generality, and there is no turnover of agents. We also consider the case with homogeneous agents having the same socialization probability and the case with heterogeneous agents having different socialization probabilities.

In the next section, we will study three cases of (i) homogeneous agents in a closed system, (ii) heterogeneous agents

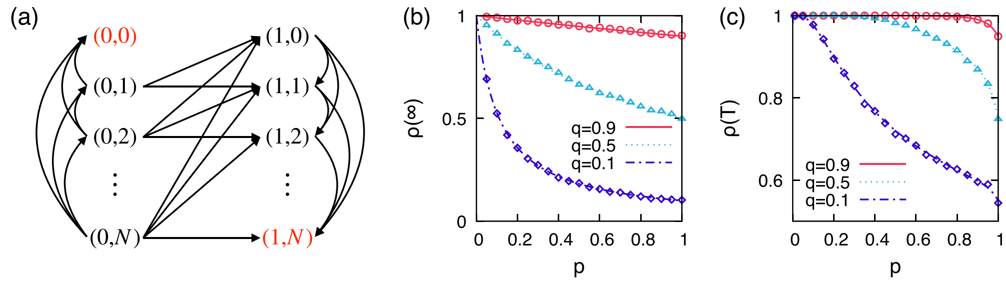

FIG. 2. (a) Transition structure between states of the system. Each state is denoted by (c, n), where c is the belief of the code and n is the number of agents whose belief is 1. Each arrow indicates a transition from one state to another, while arrows from a state to itself are omitted for better visualization. Note that both (0,0) and $(1, N)$ are absorbing states. (b) Analytic solutions of $\rho(\infty)$ in Eq. (15) (lines) with simulation results (symbols) for the case with $N = 40$ and an initial condition that $P_{cn}(0) = \delta_{c,0}\delta_{n,1}$. (c) Numerical estimation of $\rho(T)$ using the master equation in Eq. (7) (lines) for the case with $N = 40$ and an initial condition that $P_{cn}(0) = \frac{1}{2}\delta_{n,N/2}$ for each $c \in \{0, 1\}$. Here, T is the first time step satisfying $|\rho(T) - \rho(T-1)| < 10^{-4}$. The corresponding simulation results are shown with symbols. In (b) and (c), each symbol was averaged over 10^4 different runs. Standard errors are omitted as they are smaller than the symbols.

in a closed system, and (iii) homogeneous agents in an open system, which are in the same order as in March's original paper, whereas the case with heterogeneous agents in an open system is not considered in our work as in March's original paper [1].

III. RESULTS

A. Homogeneous learning in a closed system

We consider the homogeneous learning model in a closed system. That is, $r(t) = 1$ for all t, there is no turnover of agents, and every agent has the same learning probability p, i.e.,

$$p^{(i)} = p \quad \text{for} \quad i = 1, \ldots, N. \tag{3}$$

The state of the system at each time step t can be summarized in terms of the code's belief $c(t)$ and the number of agents whose belief matches the reality, which we denote by $n(t) \in \{0, \ldots, N\}$. Precisely, $n(t)$ is defined as

$$n(t) \equiv \sum_{i=1}^{N} \delta_{\sigma_i(t), r(t)}. \tag{4}$$

In our case with $r(t) = 1$, one simply has $n(t) = \sum_i \sigma_i(t)$. Then the expected density of agents with the belief matching the reality is given by

$$\rho(t) \equiv \left\langle \frac{n(t)}{N} \right\rangle, \tag{5}$$

which can be interpreted as the expected belief of a randomly chosen agent, or average individual knowledge [3].

Depending on the initial belief of the code, two scenarios are possible. First, if $c(0) = 1$, the code does not change its belief because it already coincides with the reality, and the agents' beliefs will eventually converge to the value of 1 by Eq. (1). It implies an absorbing state that the code and all agents share the same value as the reality, which is denoted by $(c, n) = (1, N)$. Second, if $c(0) = 0$, $n(t)$ will decrease until the code's belief changes to 1 by Eq. (2) as long as there is at least one agent with belief of 1. Once the code's belief becomes 1, $n(t)$ will increase to reach the absorbing state $(c, n) = (1, N)$. However, this is not always the case; $n(t)$ may reach 0 before $c(t)$ changes to 1, implying that both the code and agents have the belief of 0 without further dynamics. This indicates another absorbing state $(c, n) = (0, 0)$. Figure 2(a) shows the transition structure between states with two absorbing states emphasized in red.

For the analysis, let us denote by $P_{cn}(t)$ the probability that at time step t, the code's belief is c and there are exactly n agents with belief of 1. These probabilities satisfy the normalization condition as

$$\sum_{c=0}^{1} \sum_{n=0}^{N} P_{cn}(t) = 1. \tag{6}$$

They evolve according to the following master equation in discrete time:

$$P_{cn}(t+1) = \sum_{c'n'} W_{c'n' \to cn} P_{c'n'}(t), \tag{7}$$

where the transition probabilities read (see the Appendix)

$$\begin{aligned} W_{00\to cn} &= \delta_{0,c}\delta_{0,n}, \\ W_{0n'(\neq 0)\to 0n} &= \begin{cases} \binom{n'}{n} p^{n'-n} \bar{p}^{n} \bar{q} & \text{if } n' \geqslant n \\ 0 & \text{if } n' < n, \end{cases} \\ W_{0n'(\neq 0)\to 1n} &= \begin{cases} \binom{n'}{n} p^{n'-n} \bar{p}^{n} q & \text{if } n' \geqslant n \\ 0 & \text{if } n' < n, \end{cases} \\ W_{1n'\to 0n} &= 0 \ \forall n', n, \\ W_{1n'\to 1n} &= \begin{cases} \binom{N-n'}{N-n} p^{n-n'} \bar{p}^{N-n} & \text{if } n' \leqslant n \\ 0 & \text{if } n' > n. \end{cases} \end{aligned} \tag{8}$$

Here we have used $\bar{p} \equiv 1 - p$ and $\bar{q} \equiv 1 - q$. Calculating Eq. (7) recursively with any initial condition $\{P_{cn}(0)\}$, one can, in principle, obtain $P_{cn}(t)$ for any c, n, and t, and hence $\rho(t)$ in Eq. (5), i.e.,

$$\rho(t) = \frac{1}{N} \sum_{c=0}^{1} \sum_{n=0}^{N} n P_{cn}(t). \tag{9}$$

We focus on steady states of the model. It is obvious that all initial probabilities eventually end up with two absorbing states, i.e., $(c, n) = (0, 0)$ and $(1, N)$, implying

$P_{00}(\infty)+P_{1N}(\infty)=1$. Thus, the average individual knowledge in Eq. (9) reads

$$\rho(\infty)=P_{1N}(\infty)=1-P_{00}(\infty). \tag{10}$$

As the simplest yet nontrivial case, let us consider an initial condition that $P_{cn}(0)=\delta_{c,0}\delta_{n,1}$, namely, $P_{01}(0)=1$ and $P_{cn}(0)=0$ for all other states $(c,n)\neq(0,1)$. Using

$$W_{01\to 01}=\bar{p}\bar{q}\equiv\alpha \tag{11}$$

and

$$W_{01\to 00}=p\bar{q}\equiv\beta, \tag{12}$$

the master equations for P_{01} and P_{00} [Eq. (7)] are written as follows:

$$P_{01}(t+1)=\alpha P_{01}(t),$$
$$P_{00}(t+1)=P_{00}(t)+\beta P_{01}(t). \tag{13}$$

Master equations for all other states than $P_{01}(t)$ and $P_{00}(t)$ are irrelevant to calculate $P_{00}(\infty)$ in Eq. (10). Since $P_{01}(t)=\alpha^t$, one obtains

$$P_{00}(t)=\beta(1+\alpha+\cdots+\alpha^{t-1}), \tag{14}$$

leading to

$$\rho(\infty)=1-\frac{\beta}{1-\alpha}=\frac{q}{p+q-pq}. \tag{15}$$

This solution is not a function of N due to the choice of the initial condition that $P_{cn}(0)=\delta_{c,0}\delta_{n,1}$.

We observe that $\rho(\infty)$ in Eq. (15) is a decreasing function of p, but an increasing function of q [Fig. 2(b)], already partly implying the qualitatively similar behavior to the simulation results of the original March model, i.e., Fig. 1 in Ref. [1]. The larger q leads to the more correct belief of agents about the reality, which is easily understood by considering that q is the learning probability of the code from superior agents. On the other hand, the effect of p on $\rho(\infty)$ is not straightforward to understand. It is because the large value of p speeds up not only the probability flow to the state (0,0), but also to the state $(1,N)$. It means that the large p always helps spread the code's belief to agents, whether or not the code's belief is correct. When the code's belief is incorrect, the large p increases the amount of flow to the state (0,0) [Fig. 2(a)]. In contrast, when the code's belief is correct, the large p does not increase the amount of flow to the state $(1,N)$, but only speeds up the flow. As we focus on the steady behavior, such an asymmetric role of p leads to the decreasing behavior of $\rho(\infty)$ as a function of p. Such a behavior has been interpreted that slow socialization allows for longer exploration, resulting in better organizational learning [1].

For general initial conditions, one can estimate $\rho(T)$ for a sufficiently large T by iterating the master equation in Eq. (7) for a given initial condition $\{P_{cn}(0)\}$. For a demonstration with a system size that is neither too small nor too big, we consider a system of $N=40$ agents and the initial condition that $P_{cn}(0)=\frac{1}{2}\delta_{n,N/2}$ for each $c\in\{0,1\}$. We estimate the value of $\rho(T)$ at the first time step T when $|\rho(T)-\rho(T-1)|<10^{-4}$ is satisfied. From the results shown in Fig. 2(c), we find that $\rho(T)$ is a decreasing function of p, but an increasing function of q, showing the same tendency as the solution of $\rho(\infty)$ in Eq. (15) for the simpler initial condition.

These exact and numerical results are supported by agent-based simulations. We perform the simulations of the model using the rules in Eqs. (1) and (2) together with Eq. (3) for the system with $N=40$ agents with the mentioned initial conditions. First, the initial condition with $P_{cn}(0)=\delta_{c,0}\delta_{n,1}$ used for the analysis is realized in the simulation such that only one agent has an initial belief of 1, while all other agents and the code have the belief of 0. Second, as for the initial condition with $P_{cn}(0)=\frac{1}{2}\delta_{n,N/2}$ for each $c\in\{0,1\}$, we set $\sigma_i(0)=1$ for $i=1,\dots,20$ and $\sigma_i(0)=0$ for the rest of the agents, while the value of $c(0)$ is randomly chosen from $\{0,1\}$ with equal probabilities. Eventually, every run ends up with one of the absorbing states, implying that $n(\infty)=0,N$. For each pair of p and q, we take the average of $n(\infty)/N$ over 10^4 different runs to get the value of $\rho(\infty)$ in Eq. (5). Such averages are shown with symbols in Figs. 2(b) and 2(c), which are indeed in good agreement with the analytical and numerical solutions, respectively.

B. Heterogeneous learning in a closed system

Next, we study the heterogeneous version of the model in a closed system with $r=1$ by using two distinct values of learning probability, i.e., by setting

$$p^{(i)}=\begin{cases}p_1 & \text{for } i=1,\dots,N_1\\ p_2 & \text{for } i=N_1+1,\dots,N,\end{cases} \tag{16}$$

where $1\leqslant N_1\leqslant N-1$. Agents with the larger (smaller) learning probability among p_1 and p_2 can be called fast (slow) learners [1]. The state of the system at each time step t can be summarized in terms of the code's belief $c(t)$, the number of agents with p_1 whose belief is 1, which we denote by $n(t)\in\{0,\dots,N_1\}$, and the number of agents with p_2 whose belief is 1, which we denote by $m(t)\in\{0,\dots,N_2\}$. Here, $N_2\equiv N-N_1$. Then the expected density of agents with belief of 1 is given as

$$\rho(t)\equiv\left\langle\frac{n(t)+m(t)}{N}\right\rangle, \tag{17}$$

which can also be interpreted as the expected belief of a randomly chosen agent.

Similarly to the homogeneous version of the model, the master equation reads

$$P_{cnm}(t+1)=\sum_{c'n'm'}W_{c'n'm'\to cnm}P_{c'n'm'}(t), \tag{18}$$

where the transition probabilities are written as

$$W_{000\to cnm}=\delta_{0,c}\delta_{0,n}\delta_{0,m},$$
$$W_{0n'm'(\neq 00)\to 0nm}=\begin{cases}\binom{n'}{n}p_1^{n'-n}\bar{p}_1^{n}\binom{m'}{m}p_2^{m'-m}\bar{p}_2^{m}\bar{q} & \text{if } n'\geqslant n \text{ and } m'\geqslant m\\ 0 & \text{otherwise},\end{cases}$$

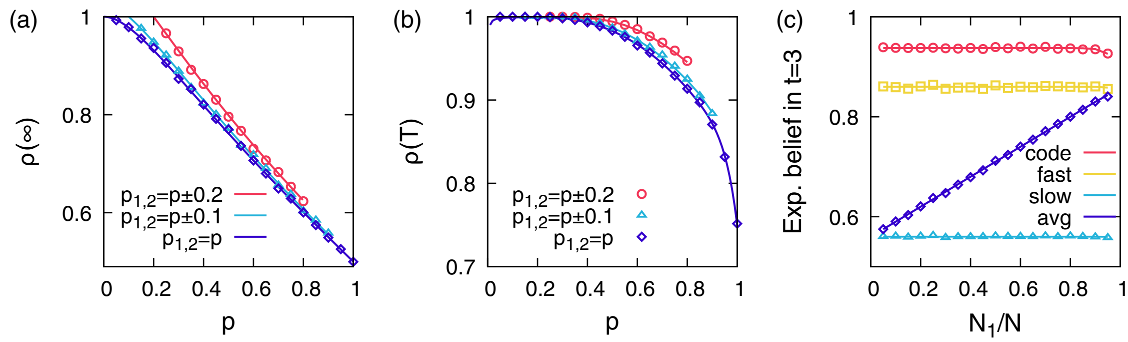

FIG. 3. (a) Analytic solutions of $\rho(\infty)$ in Eq. (23) (lines) for the case with $N = 40$ ($N_1 = N_2 = 20$) and an initial condition that $P_{cnm}(0) = \delta_{c,0}\delta_{n,1}\delta_{m,1}$. (b) Numerical estimation of $\rho(T)$ using the master equation in Eq. (18) (lines) for the case with $N = 40$ ($N_1 = N_2 = 20$) and an initial condition that $P_{cnm}(0) = \frac{1}{2}\delta_{n,N_1/2}\delta_{m,N_2/2}$ for each $c \in \{0, 1\}$. Here, T is the first time step satisfying $|\rho(T) - \rho(T-1)| < 10^{-4}$. (c) Numerical estimation of expected beliefs of the code ("code" in the figure), fast-learning agents with $p_1 = 0.9$ ("fast"), slow-learning agents with $p_2 = 0.1$ ("slow"), and all agents ("avg") at $t = 3$ using the master equation in Eq. (18) for the case with $N = 40$ and $N_1 = 2, 4, 6, \dots, 38$ (lines). We use an initial condition that $P_{cnm}(0) = \frac{1}{2}\delta_{n,N_1/2}\delta_{m,N_2/2}$ for each $c \in \{0, 1\}$ and for each N_1. In all panels, simulation results are shown with symbols, each symbol was averaged over 2×10^4 different runs, and standard errors are omitted as they are smaller than the symbols.

$$\begin{aligned}
W_{0n'm'(\neq 00)\to 1nm} &= \begin{cases} \binom{n'}{n} p_1^{n'-n}\bar{p}_1^{n}\binom{m'}{m}p_2^{m'-m}\bar{p}_2^{m} q & \text{if } n' \geqslant n \text{ and } m' \geqslant m \\ 0 & \text{otherwise,} \end{cases}\\
W_{1n'm'\to 0nm} &= 0 \;\forall n', m', n, m,\\
W_{1n'm'\to 1nm} &= \begin{cases} \binom{N_1-n'}{N_1-n} p_1^{n-n'}\bar{p}_1^{N_1-n}\binom{N_2-m'}{N_2-m}p_2^{m-m'}\bar{p}_2^{N_2-m} & \text{if } n' \leqslant n \text{ and } m' \leqslant m \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{19}$$

Here, we have used $\bar{p}_1 \equiv 1 - p_1$ and $\bar{p}_2 \equiv 1 - p_2$. It is obvious that there are two absorbing states, i.e., (0,0,0) and $(1, N_1, N_2)$, implying that $P_{000}(\infty) + P_{1N_1N_2}(\infty) = 1$. Calculating Eq. (18) recursively with any initial condition $\{P_{cnm}(0)\}$, one can, in principle, obtain $P_{cnm}(t)$ for any c, n, m, and t, and hence $\rho(t)$ in Eq. (17).

As the simplest yet nontrivial case, let us consider an initial condition that $P_{cnm}(0) = \delta_{c,0}\delta_{n,1}\delta_{m,1}$. Denoting

$$\begin{aligned}
\alpha_{i_1i_2} &\equiv W_{0i_1i_2\to 0i_1i_2},\\
\beta_{i_1i_2,j_1j_2} &\equiv W_{0i_1i_2\to 0j_1j_2}\; [(i_1, i_2) \neq (j_1, j_2)],
\end{aligned} \tag{20}$$

the master equations for P_{011}, P_{010}, P_{001}, and P_{000} [Eq. (18)] are written as follows:

$$\begin{aligned}
P_{011}(t+1) &= \alpha_{11}P_{011}(t),\\
P_{010}(t+1) &= \alpha_{10}P_{010}(t) + \beta_{11,10}P_{011}(t),\\
P_{001}(t+1) &= \alpha_{01}P_{001}(t) + \beta_{11,01}P_{011}(t),\\
P_{000}(t+1) &= P_{000}(t) + \beta_{11,00}P_{011}(t) + \beta_{10,00}P_{010}(t)\\
&\quad + \beta_{01,00}P_{001}(t),
\end{aligned} \tag{21}$$

where $\alpha_{11} = \bar{p}_1\bar{p}_2\bar{q}$, $\alpha_{10} = \bar{p}_1\bar{q}$, $\alpha_{01} = \bar{p}_2\bar{q}$, $\beta_{11,10} = \bar{p}_1p_2\bar{q}$, $\beta_{11,01} = p_1\bar{p}_2\bar{q}$, $\beta_{11,00} = p_1p_2\bar{q}$, $\beta_{10,00} = p_1\bar{q}$, and $\beta_{01,00} = p_2\bar{q}$. After some algebra, one obtains

$$\begin{aligned}
P_{000}(\infty) &= \frac{\beta_{11,00}}{1-\alpha_{11}} + \frac{\beta_{11,10}\beta_{10,00}}{(1-\alpha_{11})(1-\alpha_{10})}\\
&\quad + \frac{\beta_{11,01}\beta_{01,00}}{(1-\alpha_{11})(1-\alpha_{01})},
\end{aligned} \tag{22}$$

leading to

$$\rho(\infty) = 1 - \frac{p_1p_2\bar{q}}{1-\bar{p}_1\bar{p}_2\bar{q}}\left(1 + \frac{\bar{p}_1\bar{q}}{1-\bar{p}_1\bar{q}} + \frac{\bar{p}_2\bar{q}}{1-\bar{p}_2\bar{q}}\right). \tag{23}$$

This result is not a function of N_1 and N_2 due to the choice of the initial condition that $P_{cnm}(0) = \delta_{c,0}\delta_{n,1}\delta_{m,1}$. It is straightforward to prove that setting $p_1 = p_2 = p$ reduces the solution in Eq. (23) to the solution of the homogeneous model with the initial condition that $P_{cn}(0) = \delta_{c,0}\delta_{n,2}$.

To demonstrate the effect of heterogeneous learning on $\rho(\infty)$ in Eq. (23), we parametrize $p_1 = p + \delta$ and $p_2 = p - \delta$ with non-negative δ and $p \in (\delta, 1-\delta]$. Here, δ controls the degree of heterogeneity of the agents. As shown in Fig. 3(a), the larger δ leads to the higher values of $\rho(\infty)$ in Eq. (23) for the entire range of p, which is consistent with the simulation results of the original March model, i.e., Fig. 2 in Ref. [1]. Such behaviors can be essentially understood by comparing the transition probability $W_{011\to 000} = p_1p_2\bar{q}$ in the heterogeneous model to its counterpart $W_{02\to 00} = p^2\bar{q}$ in the homogeneous model [Eq. (8)] to get

$$\frac{W_{011\to 000}}{W_{02\to 00}} = 1 - \frac{\delta^2}{p^2} \leqslant 1. \tag{24}$$

It implies that for positive δ, the probability flow to the absorbing state (0,0,0) in the heterogeneous model is always smaller than the flow to the absorbing state (0,0) in the homogeneous model, and hence the larger $\rho(\infty)$ for the heterogeneous model than for the homogeneous model. We also remark that the ratio in Eq. (24) gets closer to 1 for the larger value of p, and hence the smaller gap between the heterogeneous and homogeneous models. Such an expectation is indeed the case, as depicted in Fig. 3(a).

For general initial conditions, we numerically estimate $\rho(T)$ for a sufficiently large T by iterating the master equation in Eq. (18) for a given initial condition $\{P_{cnm}(0)\}$. For a demonstration, we consider a system of $N = 40$ agents ($N_1 = N_2 = 20$) and the initial condition that $P_{cnm}(0) = \frac{1}{2}\delta_{n,N_1/2}\delta_{m,N_2/2}$ for each $c \in \{0, 1\}$. We estimate the value of $\rho(T)$ at the first time step T when $|\rho(T) - \rho(T-1)| < 10^{-4}$ is satisfied. From the results shown in Fig. 3(b), we find that $\rho(T)$ has higher values for more heterogeneous systems.

We also perform the agent-based simulations of the heterogeneous model for the system with $N = 40$ agents ($N_1 = N_2 = 20$) with the mentioned initial conditions. First, the initial condition with $P_{cnm}(0) = \delta_{c,0}\delta_{n,1}\delta_{m,1}$ is realized in the simulation such that one agent with p_1 and one agent with p_2 have an initial belief of 1, while all other agents as well as the code have the belief of 0. Second, as for the initial condition with $P_{cnm}(0) = \frac{1}{2}\delta_{n,N_1/2}\delta_{m,N_2/2}$ for each $c \in \{0, 1\}$, we set $\sigma_i(0) = 1$ for $i = 1, \dots, 10, 21, \dots, 30$ and $\sigma_i(0) = 0$ for the rest of the agents, while the value of $c(0)$ is randomly chosen from $\{0, 1\}$ with equal probabilities. Eventually, every run ends up with one of the absorbing states, implying that $n(\infty) + m(\infty) = 0, N$. For each combination of p_1, p_2, and q, we take the average of $[n(\infty) + m(\infty)]/N$ over 2×10^4 different runs to get the value of $\rho(\infty)$ in Eq. (17). Such averages are shown with symbols in Figs. 3(a) and 3(b), which are indeed in good agreement with analytical and numerical solutions, respectively.

Finally, we note that our setup for heterogeneous agents is different from that in the original March model [1]. In the original paper, the heterogeneity was controlled by the number of agents having p_1, i.e., N_1, while learning probabilities were fixed to be $p_1 = 0.9$ and $p_2 = 0.1$. We test such original setup using our simplified model both by estimating $\rho(3)$ from the master equations in Eq. (18) and by performing agent-based simulations up to $t = 3$. For a system of $N = 40$ agents, we consider $N_1 = 2, 4, 6, \dots, 38$ with the initial condition that $P_{cnm}(0) = \frac{1}{2}\delta_{n,N_1/2}\delta_{m,N_2/2}$ for each $c \in \{0, 1\}$. In addition to the expected belief of all agents in Eq. (17), we measure the expected belief of fast-learning agents with p_1, that of slow-learning agents with p_2, and that of the code. Results from the numerical estimation of master equations and from agent-based simulations are in good agreement with each other, as depicted in Fig. 3(c). These results show the qualitatively same behaviors as in the original March model, i.e., Fig. 3 in Ref. [1].

C. Homogeneous learning in an open system

Finally, we study the effects of turnover of agents and turbulence of the external reality on organizational learning. For this, we focus on the simplest yet nontrivial case with $N = 1$, indicating that there is only one agent in the system. This agent's belief is denoted by $\sigma(t)$. The case with general $N > 1$ can be studied too within our framework.

We first consider the system with turnover of the agent only, while $r(t) = 1$ for all t, namely, $u > 0$ and $v = 0$. Let us denote by $P_{c\sigma}(t)$ the probability that at time step t, the code's belief is c and the agent's belief is σ. These probabilities satisfy the normalization condition as

$$\sum_{c,\sigma\in\{0,1\}} P_{c\sigma}(t) = 1. \quad (25)$$

They evolve according to the following master equation in discrete time:

$$P_{c\sigma}(t+1) = \sum_{c'\sigma'} W_{c'\sigma'\to c\sigma} P_{c'\sigma'}(t), \quad (26)$$

where the transition probabilities, using $u' \equiv u/2$ and $\bar{u}' \equiv 1 - u'$, are as follows:

$$\begin{aligned} &W_{00\to00} = \bar{u}',\ W_{00\to01} = u',\\ &W_{01\to00} = (p\bar{u}' + \bar{p}u')\bar{q},\ W_{01\to01} = (\bar{p}\bar{u}' + pu')\bar{q},\\ &W_{01\to10} = (p\bar{u}' + \bar{p}u')q,\ W_{01\to11} = (\bar{p}\bar{u}' + pu')q,\\ &W_{10\to10} = \bar{p}\bar{u}' + pu',\ W_{10\to11} = p\bar{u}' + \bar{p}u',\\ &W_{11\to10} = u',\ W_{11\to11} = \bar{u}', \end{aligned} \quad (27)$$

and all other transition probabilities are zero. Note that due to $u > 0$, both $(c, \sigma) = (0, 0)$ and (1,1) are no longer absorbing states.

For the steady state, we derive the analytical solution of $\rho(\infty)$ as

$$\rho(\infty) = \sum_{c\in\{0,1\}} P_{c1}(\infty) = \frac{p + \left(\frac{1}{2} - p\right)u}{p + (1-p)u}. \quad (28)$$

This solution is independent of the initial condition and it is not a function of q because the change of the code's belief from 0 to 1 is irreversible; as long as $q > 0$, all initial probabilities end up with states with $c = 1$. We also find that $\lim_{u\to0}\rho(\infty) = 1$ and $\rho(\infty) = 1/2$ for $u = 1$, both of which are irrespective of p. That is, $\rho(\infty)$ is a decreasing function of u for $u > 0$, whereas for $u = 0$, it can have the finite value of less than one, e.g., given in Eq. (15), as long as $p > 0$. Thus one can conclude that $\rho(\infty)$ shows an "increasing" and then decreasing behavior in the range of $0 \leqslant u \leqslant 1$. This argument is important to discuss about the optimal turnover of the agent that maximizes the effectiveness of organizational learning.

Next, we focus on the transient dynamics instead of the steady state. Starting from the initial condition that $P_{c\sigma}(0) = \delta_{c,0}\delta_{\sigma,1}$, we obtain, e.g., at $t = 2$,

$$\begin{aligned} \rho(2) =& \left(\tfrac{1}{2} - p\right)(1-p)u^2 - \left(2p^2 + pq - \tfrac{7}{2}p + 1\right)u\\ &+ (1-p)^2 + pq. \end{aligned} \quad (29)$$

It turns out that $\rho(2)$ is a quadratic function of u, meaning that it can have either a maximum or minimum value for the range of $u \in (0, 1)$, or it may be a monotonically increasing or decreasing function of u, depending on the choice of p and q. We also observe that $\rho(t)$ is, in general, the tth degree polynomial function of u that may show more complex behaviors. Figure 4(a) summarizes various behaviors of $\rho(2)$ in the plane of (p, q), and Fig. 4(b) depicts $\rho(2)$ as a function of u for several cases of p and q. For example, for sufficiently large q, $\rho(2)$ is a monotonically decreasing function of u, irrespective of p. It implies that if the code learns fast from the superior agent, the maximal organizational learning is achieved when

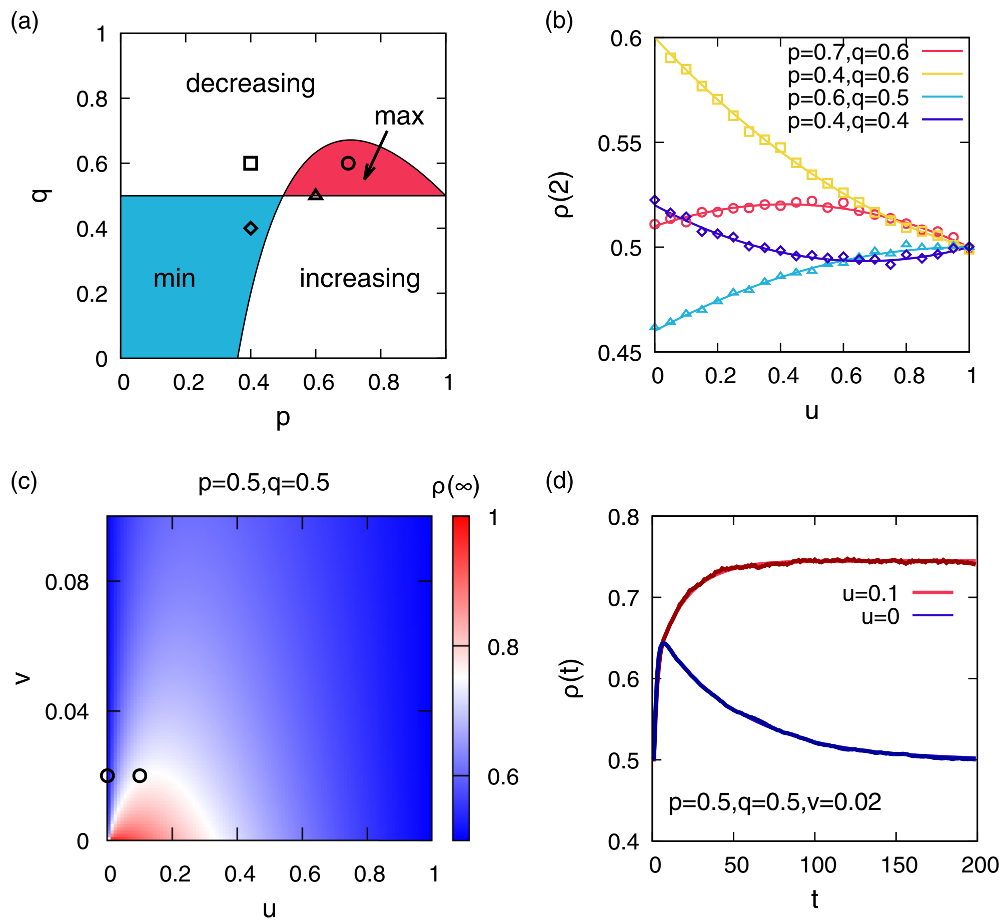

FIG. 4. (a) Different behaviors of the analytic solution of $\rho(2)$ in Eq. (29) as a function of u depicted in the plane of (p, q). $\rho(2)$ can have either a maximum value for $u \in (0, 1)$ (red shade, denoted by "max") or a minimum value for $u \in (0, 1)$ (blue shade, denoted by "min"), or it can monotonically increase ("increasing") or decrease ("decreasing"). Four empty symbols are chosen to demonstrate their different functional forms of $\rho(2)$ in (b). (b) Analytic solutions of $\rho(2)$ in Eq. (29) as a function of u for several combinations of p and q (lines) with corresponding simulation results (symbols). We use an initial condition that $P_{c\sigma}(0) = \delta_{c,0}\delta_{\sigma,1}$. (c) Heat maps of the analytic solution of $\rho(\infty)$ in Eq. (35) as a function of u and v for the case with $p = q = 0.5$. Two empty symbols are chosen to demonstrate different behaviors of $\rho(t)$ in (d). (d) Numerical estimation of $\rho(t)$ using the master equation (solid lines) and simulation results (dotted lines) for $u = 0$ and 0.1 when $p = 0.5$, $q = 0.5$, and $v = 0.02$. In (b) and (d), each symbol was averaged over 2×10^5 different runs. Standard errors are omitted as they are smaller than the symbols.

there is no turnover. This result can be understood by considering the fact that the turnover introduces randomness or new information from outside of the system. In contrast, if the code learns slowly from the superior agent but the agent learns fast from the code, the maximal organizational learning is achieved for the largest turnover. In such case, without turnover, both the code and agent are likely to be stuck in a suboptimal situation. Thus, the strong turnover may help the system to evade it. Precisely, we find the increasing and then decreasing behavior of $\rho(2)$ for $p = 0.7$ and $q = 0.6$, and the monotonically decreasing behavior of $\rho(2)$ for $p = 0.4$ and $q = 0.6$, which are consistent with the results of the original March model, e.g., Fig. 4 in Ref. [1].

We now consider the effect of turbulence on the organizational learning in the presence of turnover of the agent. For this, we define an extended system consisting of both the system and the reality, whose states can be denoted by $(r, c, \sigma) \in \{0, 1\}^3$. Let us denote by $P_{rc\sigma}(t)$ the probability that at time step t the reality is r, the code's belief is c, and the agent's belief is σ. These probabilities satisfy the normalization condition as

$$\sum_{r,c,\sigma\in\{0,1\}} P_{rc\sigma}(t) = 1. \tag{30}$$

They evolve according to the following master equation in discrete time:

$$P_{rc\sigma}(t+1) = \sum_{r'c'\sigma'} W_{r'c'\sigma'\to rc\sigma} P_{r'c'\sigma'}(t). \tag{31}$$

Denoting $v' \equiv v/2$ and $\bar{v}' \equiv 1 - v'$ and using Eq. (27), we get the transition probabilities for Eq. (31) as follows:

$$\begin{aligned} W_{0c'\sigma'\to 0c\sigma} &= \overline{W}_{c'\sigma'\to c\sigma}\,\bar{v}', \\ W_{0c'\sigma'\to 1c\sigma} &= \overline{W}_{c'\sigma'\to c\sigma}\,v', \\ W_{1c'\sigma'\to 0c\sigma} &= W_{c'\sigma'\to c\sigma}\,v', \\ W_{1c'\sigma'\to 1c\sigma} &= W_{c'\sigma'\to c\sigma}\,\bar{v}', \end{aligned} \tag{32}$$

where we have used

$$\overline{W}_{c'\sigma'\to c\sigma} \equiv W_{1-c',1-\sigma'\to 1-c,1-\sigma}. \tag{33}$$

As $r(t)$ is no longer constant, the average individual knowledge is obtained as

$$\rho(t) = \sum_{r,c,\sigma\in\{0,1\}} \delta_{r,\sigma} P_{rc\sigma}(t). \tag{34}$$

After some algebra, we derive an exact solution of $\rho(\infty)$ for the steady state as follows:

$$\begin{aligned}\rho(\infty) = \big[&\big(p^2q - p^2 - \tfrac{5}{2}pq + 2p + q - 1\big)u^2v^2 - \big(p^2q + p^2 - \tfrac{7}{2}pq + 2p + \tfrac{3}{2}q - 1\big)u^2v - \big(p - \tfrac{1}{2}\big)qu^2\\ &- \big(2p^2q - 2p^2 - \tfrac{9}{2}pq + 3p + \tfrac{3}{2}q - 1\big)uv^2 + \big(2p^2q - 2p^2 - \tfrac{9}{2}pq + 2p + q\big)uv + pqu\\ &+ \big(p^2q - p^2 - 2pq + p + \tfrac{1}{2}q\big)v^2 - p(pq - p - q)v\big]\big/[2(p^2q - p^2 - 2pq + 2p + q - 1)u^2v^2\\ &- (2p^2q - 2p^2 - 5pq + 4p + 3q - 2)u^2v + (1 - p)qu^2 - (4p^2q - 4p^2 - 8pq + 6p + 3q - 2)uv^2\\ &+ (4p^2q - 4p^2 - 7pq + 4p + 2q)uv + pqu + (2p^2q - 2p^2 - 4pq + 2p + q)v^2 - 2p(pq - p - q)v].\end{aligned} \tag{35}$$

This analytical solution is depicted as a heat map in Fig. 4(c) for the case with $p = q = 0.5$. We find that for each value of turbulence v, there exists an optimal turnover probability $u^* \in (0, 1)$ that maximizes the effectiveness of organizational learning. Such an optimal turnover probability for a given v is obtained as

$$u^*(v) = \frac{v^2 + 3v - \sqrt{2v(v+3)(3v+1)}}{v^2 - 3v - 2}, \tag{36}$$

which is an increasing function of v. It implies that the system is required to have the larger turnover to adapt to the more turbulent external reality. Yet the value of $\rho(\infty)$ with the optimal turnover tends to decrease with v [Fig. 4(c)]. We remind the reader that smaller $\rho(\infty)$ means more incorrect belief of the agent in the organization, and hence less effective organizational learning.

Finally, we look at the transient dynamics of $\rho(t)$ for different values of u when p, q, and v are given. We numerically obtain $\rho(t)$ by iterating the master equation in Eq. (26) using the initial condition that $P_{rc\sigma(0)} = 1/8$ for each state. Numerical results are depicted as solid lines in Fig. 4(d). The agent-based simulations are also performed using the initial condition that each of $r(0)$, $c(0)$, and $\sigma(0)$ is randomly and independently drawn from $\{0, 1\}$. Simulation results are shown as dotted lines in Fig. 4(d), which are in good agreement with the numerical results. These results are also qualitatively similar to those in the original March model, e.g., Fig. 5 in Ref. [1].

IV. CONCLUSION

In our work, the celebrated organizational learning model proposed by March [1] has been simplified, enabling us to explicitly write the master equation for the dynamics of the model. We have successfully derived exact solutions for the simplest yet nontrivial cases and numerically estimated quantities of interest using the master equations; both results are found to be in good agreement with agent-based simulation results. Our results help to provide a rigorous understanding not only of the simplified model, but also of the original March model, to a large extent.

Our theoretical framework for the simplified March model can be applied to the original March model, as well as variants of the March's model that incorporate other relevant factors such as forgetting of the beliefs [3,11] and direct interaction and communication between agents in the organization [4–6]. For modeling the interaction structure between agents, various network models might be deployed [25,29–31]. In conclusion, we expect to gain deeper insights into the organizational learning using our analytical approach.

ACKNOWLEDGMENTS

H.-H.J. acknowledges financial support by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant No. 2022R1A2C1007358).

APPENDIX: DERIVATION OF TRANSITION PROBABILITIES IN EQ. (8)

By definition of the model, it is obvious that $W_{00\to cn} = \delta_{0,c}\delta_{0,n}$ and $W_{1n'\to 0n} = 0\ \forall n', n$. Here, we derive other transition probabilities. The transition from a state $(0, n')$ with $n' > 0$ to the state $(0, n)$ occurs in the case when c keeps having the value of 0 and when $n' \geqslant n$ because only agents with belief 1 may change their beliefs to 0 due to the socialization from the code. The probability that c keeps having the value of 0 is $\bar{q}$, while the probability that n' reduces to n is $\binom{n'}{n}p^{n'-n}\bar{p}^n$. Since these two events are independent of each other, we get $W_{0n'(\neq 0)\to 0n}$ as a multiplication of their corresponding probabilities:

$$W_{0n'(\neq 0)\to 0n} = \begin{cases}\binom{n'}{n}p^{n'-n}\bar{p}^n\bar{q} & \text{if } n' \geqslant n\\ 0 & \text{if } n' < n.\end{cases} \tag{A1}$$

As mentioned, since n cannot be bigger than n', we have $W_{0n'(\neq 0)\to 0n} = 0$ for $n' < n$. Similarly, we get $W_{0n'(\neq 0)\to 1n}$ by replacing $\bar{q}$ by q in $W_{0n'(\neq 0)\to 0n}$ because the probability that c changes its value to 1 is q,

$$W_{0n'(\neq 0)\to 1n} = \begin{cases}\binom{n'}{n}p^{n'-n}\bar{p}^n q & \text{if } n' \geqslant n\\ 0 & \text{if } n' < n.\end{cases} \tag{A2}$$

Finally, if $c = 1$ already, it does not change its value. Only agents with belief 0 may change their beliefs to 1 due to the socialization from the code, implying that the number of agents whose belief is 1 cannot decrease, i.e., $n' \leqslant n$. The probability that n' increases to n is $\binom{N-n'}{N-n} p^{n-n'} \bar{p}^{N-n}$. Thus, one gets

$$W_{1n' \to 1n} = \begin{cases} \binom{N-n'}{N-n} p^{n-n'} \bar{p}^{N-n} & \text{if } n' \leqslant n \\ 0 & \text{if } n' > n. \end{cases} \tag{A3}$$

[1] J. G. March, Exploration and exploitation in organizational learning, Org. Sci. **2**, 71 (1991).
[2] S. Rodan, Exploration and exploitation revisited: Extending March's model of mutual learning, Scand. J. Manag. **21**, 407 (2005).
[3] S. Blaschke and D. Schoeneborn, The forgotten function of forgetting: Revisiting exploration and exploitation in organizational learning, Soz. Syst. **12**, 100 (2006).
[4] K. D. Miller, M. Zhao, and R. J. Calantone, Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model, Acad. Manag. J. **49**, 709 (2006).
[5] G. C. Kane and M. Alavi, Information technology and organizational learning: An investigation of exploration and exploitation processes, Org. Sci. **18**, 796 (2007).
[6] T. Kim and M. Rhee, Exploration and exploitation: internal variety and environmental dynamism, Strateg. Org. **7**, 11 (2009).
[7] C. Fang, J. Lee, and M. A. Schilling, Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning, Org. Sci. **21**, 625 (2010).
[8] M. Sachdeva, Encounter with March's organizational learning model, Rev. Integrat. Bus. Econ. Res. **2**, 602 (2013).
[9] M. A. Schilling and C. Fang, When hubs forget, lie, and play favorites: Interpersonal network structure, information distortion, and organizational learning, Strateg. Manag. J. **35**, 974 (2014).
[10] S. S. Chanda and S. Ray, Optimal exploration and exploitation: The managerial intentionality perspective, Comput. Math. Org. Theory **21**, 247 (2015).
[11] K. D. Miller and D. Martignoni, Organizational learning with forgetting: Reconsidering the exploration–exploitation tradeoff, Strateg. Org. **14**, 53 (2016).
[12] S. S. Chanda, Inferring final organizational outcomes from intermediate outcomes of exploration and exploitation: the complexity link, Comput. Math. Org. Theory **23**, 61 (2017).
[13] S. S. Chanda, S. Ray, and B. Mckelvey, The continuum conception of exploration and exploitation: An update to March's theory, Management **21**, 1050 (2018).
[14] S. S. Chanda and K. D. Miller, Replicating agent-based models: Revisiting March's exploration–exploitation study, Strateg. Org. **17**, 425 (2019).
[15] D. A. Marín-Idárraga, J. M. Hurtado González, and C. Cabello Medina, Factors affecting the effect of exploitation and exploration on performance: A meta-analysis, BRQ Bus. Res. Quart. **25**, 312 (2022).
[16] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, Rev. Mod. Phys. **81**, 591 (2009).
[17] D. Acemoglu and A. Ozdaglar, Opinion dynamics and learning in social networks, Dynamic Games Applic. **1**, 3 (2011).
[18] P. Sen and B. K. Chakrabarti, *Sociophysics: An Introduction* (Oxford University Press, Oxford, 2014).
[19] A. Sîrbu, V. Loreto, V. D. P. Servedio, and F. Tria, Opinion dynamics: Models, extensions and external effects, in *Participatory Sensing, Opinions and Collective Awareness*, edited by V. Loreto, M. Haklay, A. Hotho, V. D. Servedio, G. Stumme, J. Theunis, and F. Tria (Springer International, Cham, 2017), pp. 363–401.
[20] A. V. Proskurnikov and R. Tempo, A tutorial on modeling and analysis of dynamic social networks. Part I, Annu. Rev. Control **43**, 65 (2017).
[21] A. V. Proskurnikov and R. Tempo, A tutorial on modeling and analysis of dynamic social networks. Part II, Annu. Rev. Control **45**, 166 (2018).
[22] A. Baronchelli, The emergence of consensus: a primer, R. Soc. Open Sci. **5**, 172189 (2018).
[23] B. D. O. Anderson and M. Ye, Recent advances in the modeling and analysis of opinion dynamics on influence networks, Intl. J. Automat. Comput. **16**, 129 (2019).
[24] H. Noorazar, Recent advances in opinion propagation dynamics: A 2020 survey, Europ. Phys. J. Plus **135**, 521 (2020).
[25] A.-L. Barabási and M. Pósfai, *Network Science* (Cambridge University Press, Cambridge, 2016).
[26] M. Laguna, S. Risau Gusman, G. Abramson, S. Gonçalves, and J. Iglesias, The dynamics of opinion in hierarchical organizations, Physica A **351**, 580 (2005).
[27] S.-G. Han, J. Um, and B. J. Kim, Voter model on a directed network: Role of bidirectional opinion exchanges, Phys. Rev. E **81**, 057103 (2010).
[28] E. Lee, P. Holme, and S. H. Lee, Modeling the dynamics of dissent, Physica A **486**, 262 (2017).
[29] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, Network analysis in the social sciences, Science **323**, 892 (2009).
[30] M. E. J. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford, 2018).
[31] F. Menczer, S. Fortunato, and C. A. Davis, *A First Course in Network Science* (Cambridge University Press, Cambridge, 2020).