SUPPLEMENTAL MATERIAL FOR "ANOMALY, CLASS DIVISION, AND DECOUPLING IN INCOME DYNAMICS"

Jaeseok Hur,[1] Meesoon Ha,[2,*] and Hawoong Jeong[1,3,†]

[1] *Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea*
[2] *Department of Physics Education, Chosun University, Gwangju 61452, Korea*
[3] *Center of Complex Systems, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea*

(Dated: March 23, 2026)

CONTENTS

* Corresponding author; msha@chosun.ac.kr
† Corresponding author; hjeong@kaist.edu

I. BM MODEL ON RING TOPOLOGIES

This section is the extended version of *Appendix A* in **End Matter** (EM), where we provide all the details (analytical derivations and numerical confirmations) for the homogeneous Bouchaud-Mézard (BM) model on ring topologies.

A. Variance, Covariance, and Field exponent η

For a 1D ring topology, Eq. (A2) in EM can be rewritten as follows:

$$dc_n = J\left(\frac{c_{n-1}+c_{n+1}}{2} - c_n\right)dt + \beta c_n dW_{t,n}, \tag{S1}$$

where $n = 0, 1, \ldots, N-1$. In the periodic boundary conditions, $c_{-1} = c_{N-1}$ and $c_N = c_0$. Let $X_n = \ln c_n + \beta^2 t/2$ and apply the Itô's lemma into Eq. (S1), it becomes

$$dX_n = d\ln c_n + \frac{1}{2}\beta^2 dt = J\left(\frac{c_{n-1}+c_{n+1}}{2c_n} - 1\right)dt + \beta dW_{t,n} - \frac{1}{2}\beta^2 dt + \frac{1}{2}\beta^2 dt. \tag{S2}$$

Using $c_n = \exp\left(X_n - \beta^2 t/2\right)$, the equation for X_n can be rewritten as follows:

$$dX_n = J\left(\frac{1}{2}e^{X_{n-1}-X_n} + \frac{1}{2}e^{X_{n+1}-X_n} - 1\right)dt + \beta dW_{t,n}. \tag{S3}$$

We assume that for a small β condition, the difference of X between neighboring nodes becomes sufficiently small. As a result, the first-order approximation should be:

$$dX_n = J\left(\frac{X_{n-1}+X_{n+1}}{2} - X_n\right)dt + \beta dW_{t,n}. \tag{S4}$$

To solve the stochastic differential equations (SDE) of Eq. (S4), we perform a discrete Fourier transform:

$$\hat{X}_{\mathrm{k}} = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} X_n e^{-i2\pi\frac{\mathrm{k}}{N}n}, \quad X_n = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\hat{X}_{\mathrm{k}} e^{i2\pi\frac{\mathrm{k}}{N}n}. \tag{S5}$$

Substituting the discrete inverse Fourier transform into Eq. (S4), it becomes

$$d\hat{X}_{\mathrm{k}}(t) = -J\left(1 - \cos\frac{2\pi\mathrm{k}}{N}\right)\hat{X}_{\mathrm{k}}(t)dt + \beta dW_{\mathrm{k}}(t), \tag{S6}$$

which is the SDE corresponding to the k-th Fourier mode. Here we can see that $\hat{X}_{\mathrm{k}}(t)$ obeys an independent *Ornstein-Uhlenbeck* (OU) process at the mode k. The solution of the OU process is

$$\hat{X}_{\mathrm{k}}(t) = \hat{X}_{\mathrm{k}}(0)e^{-Jt(1-\cos 2\pi\mathrm{k}/N)} + \beta\int_0^t e^{-J(1-\cos 2\pi\mathrm{k}/N)(t-s)}dW_{\mathrm{k}}(s) \tag{S7}$$

According to the Parseval's theorem, $\sum_{n=0}^{N-1}|X_n|^2 = \sum_{k=0}^{N-1}|\hat{X}_{\mathrm{k}}|^2$, the variance of X_n is as follows:

$$\mathrm{Var}(X_n) = \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[|X_n|^2] - \left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[X_n]\right)^2 = \frac{1}{N}\sum_{\mathrm{k}=0}^{N-1}\mathbb{E}[|\hat{X}_{\mathrm{k}}|^2] - \left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[X_n]\right)^2, \tag{S8}$$

where $\mathbb{E}[\cdot]$ is the expectation for the stochastic process. We consider the initial condition $X_n(0) = \mathrm{const}$, where all nodes have the same value of X. From the given SDE, $d\left(\sum_{n=0}^{N-1}X_n\right) = 0$, so that the mean value of X_n is always a constant, regardless of t. For this case, $\hat{X}_{\mathrm{k}}(0) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}X_n(0)e^{-i2\pi\frac{\mathrm{k}}{N}n} = \sqrt{N}X_n(0)\delta_{0\mathrm{k}}$.

To get $\mathrm{Var}(X_n)$, we calculate $\mathbb{E}[|\hat{X}_\mathrm{k}|^2]$:

$$\begin{aligned}\mathbb{E}[|\hat{X}_\mathrm{k}|^2] &= \hat{X}_\mathrm{k}^2(0)e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)} + 2\beta\hat{X}_\mathrm{k}(0)e^{-Jt(1-\cos 2\pi\mathrm{k}/N)}\int_0^t e^{-J(1-\cos 2\pi\mathrm{k}/N)(t-s)}\langle dW_\mathrm{k}(s)\rangle \\ &+ \beta^2\int_0^t\int_0^t e^{-2J(1-\cos 2\pi\mathrm{k}/N)(t-\frac{s+s'}{2})}\langle dW_\mathrm{k}(s)dW_\mathrm{k}(s')\rangle \\ &= \hat{X}_\mathrm{k}^2(0)e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)} + \beta^2\int_0^t e^{-2J(1-\cos 2\pi\mathrm{k}/N)(t-s)}ds \\ &= NX_n^2(0)\delta_{0\mathrm{k}}e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)} + \beta^2\left[\frac{1-e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)}}{2J(1-\cos 2\pi\mathrm{k}/N)}\right], \end{aligned} \tag{S9}$$

where we use $\langle dW_\mathrm{k}(s)\rangle = 0$, $\langle dW_\mathrm{k}(s)dW_\mathrm{k}(s')\rangle = \delta(s-s')ds$, and $\delta_{0\mathrm{k}}^2 = \delta_{0\mathrm{k}}$. As a result, we get

$$\begin{aligned}\mathrm{Var}(X_n) &= \left[\frac{1}{N}\sum_{\mathrm{k}=0}^{N-1} NX_n(0)^2\delta_{0\mathrm{k}}e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)} + \frac{1}{N}\sum_{\mathrm{k}=0}^{N-1}\beta^2\left[\frac{1-e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)}}{2J(1-\cos 2\pi\mathrm{k}/N)}\right]\right] - \left[X_n(0)^2\right] \\ &= \frac{1}{N}\sum_{\mathrm{k}=0}^{N-1}\beta^2\left[\frac{1-e^{-2Jt(1-\cos 2\pi\mathrm{k}/N)}}{2J(1-\cos 2\pi\mathrm{k}/N)}\right] \approx \frac{\beta^2}{4\pi J}\int_0^{2\pi}\frac{1-e^{-2Jt(1-\cos u)}}{1-\cos u}du = \frac{\beta^2}{2\pi J}\int_0^{\pi}\frac{1-e^{-2Jt(1-\cos u)}}{1-\cos u}du, \end{aligned} \tag{S10}$$

which is the result of transforming the discrete sum of Fourier mode k into a definite integral for $N\to\infty$, and the periodic property of the integrand allows us to reduce the integration interval $[0, 2\pi]$ to $[0, \pi]$. If we differentiate the expression of $\mathrm{Var}(X_n)$ with respect to t,

$$\frac{d}{dt}\mathrm{Var}(X_n) = \beta^2 e^{-2Jt}\cdot\frac{1}{\pi}\int_0^{\pi} e^{2Jt\cos u}du = \beta^2 e^{-2Jt}I_0(2Jt), \tag{S11}$$

where $I_\ell(z)$ is the modified Bessel function of the first kind, represented by

$$I_\ell(z) = \frac{1}{\pi}\int_0^{\pi} e^{z\cos\theta}\cos\ell\theta d\theta - \frac{\sin\ell\pi}{\pi}\int_0^{\infty} e^{-z\cosh q-\ell q}dq. \tag{S12}$$

We get $\mathrm{Var}(X_n)$ by solving Eq. (S11). The right-hand side function has a well-defined indefinite integral:

$$\int e^{-2Jt}I_0(2Jt)dt = te^{-2Jt}\left[I_0(2Jt) + I_1(2Jt)\right] + \mathrm{C}. \tag{S13}$$

According to initial condition $X_n(0) = \mathrm{const}$ at $t = 0$, $\mathrm{Var}(X_n(0)) = 0$ and $\lim_{t\to 0} te^{-2Jt}\left[I_0(2Jt) + I_1(2Jt)\right] = 0$, and a constant of integration $\mathrm{C} = 0$. Therefore,

$$\mathrm{Var}(X_n) = \beta^2 te^{-2Jt}[I_0(2Jt) + I_1(2Jt)]. \tag{S14}$$

The Taylor expansions of Eq. (S14) for small t and large t, respectively, are

$$\begin{cases}\beta^2 t\left(1 - \frac{(2Jt)}{2} + \frac{(2Jt)^2}{4} - \frac{5(2Jt)^3}{48} + \cdots\right) & \text{for small } t, \\ \beta^2 t\left(\sqrt{\frac{2}{\pi}}(2Jt)^{-1/2} - \frac{1}{4\sqrt{2\pi}}(2Jt)^{-3/2} - \frac{3}{64\sqrt{2\pi}}(2Jt)^{-5/2} - \cdots\right) & \text{for large } t, \end{cases} \tag{S15}$$

where the leading order terms are $\beta^2 t$ and $\beta^2\sqrt{t}/(2Ja_0)$, respectively.

With $a_0(J) = \sqrt{\pi/(4J)}$, we finally get

$$\mathrm{Var}(X_n) = \begin{cases}\beta^2 t & \text{for small } t, \\ \beta^2\sqrt{t}/(2Ja_0) & \text{for large } t. \end{cases} \tag{S16}$$

Thus, a given stochastic process in Eq. (S4) starts with normal diffusion and ends with sub-diffusion. For large t, the analytical results of $\mathrm{Var}(X_n) \sim \sqrt{t}$ is consistent with the EFT and numerical simulations.

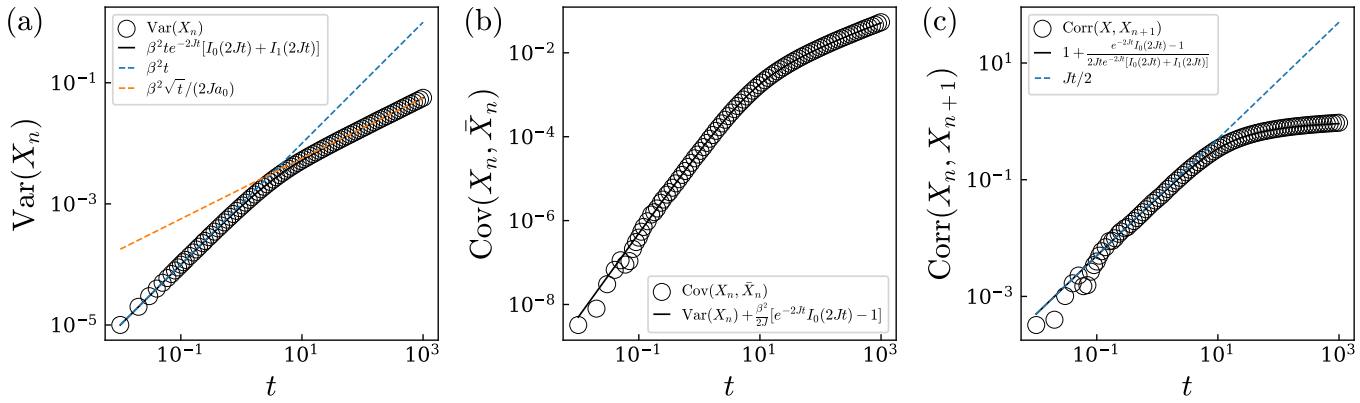

FIG. S1. (a) Variance $\mathrm{Var}(X_n)$, (b) covariance $\mathrm{Cov}(X_n,\bar{X}_n)$, and (c) correlation $\mathrm{Corr}(X_n,X_{n+1})$ in the BM model for a 1D ring topology. Open symbols ($\circ$) represent numerical simulation results and and solid lines represent the theoretical prediction. In (a), the blue dashed line $\{\beta^2 t, Jt/2\}$ is the predictions for small t and othe range dashed line $\{\beta^2\sqrt{t}/(2Ja_0)\}$ is the prediction for large t. For all cases, $N=10^4, \alpha=10^{-2}, \beta^2=10^{-3}, J=10^{-1}$, and all data are averaged over 128 ensembles.

To obtain the field exponent η related to the diffusion nature, we require the calculation of $\mathrm{Cov}(X_n,\bar{X}_n)$. First, the discrete Fourier transform of $\bar{X}_n$ is:

$$\begin{aligned}
\hat{\bar{X}}_{\mathrm{k}} &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\bar{X}_n e^{-i2\pi\frac{\mathrm{k}}{N}n} = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\frac{X_{n-1}+X_{n+1}}{2}e^{-i2\pi\frac{\mathrm{k}}{N}n} \\
&= \frac{1}{2\sqrt{N}}\sum_{n=0}^{N-1}\left[\frac{1}{\sqrt{N}}\sum_{\mathrm{k}'=0}^{N-1}(\hat{X}_{\mathrm{k}'}e^{i2\pi\frac{\mathrm{k}'}{N}(n-1)}+\hat{X}_{\mathrm{k}'}e^{i2\pi\frac{\mathrm{k}'}{N}(n+1)})\right]e^{-i2\pi\frac{\mathrm{k}}{N}n} \\
&= \frac{1}{2\sqrt{N}}\sum_{n=0}^{N-1}\left[\frac{1}{\sqrt{N}}\sum_{\mathrm{k}'=0}^{N-1}\hat{X}_{\mathrm{k}'}e^{i2\pi\frac{\mathrm{k}'}{N}n}(e^{-i2\pi\frac{\mathrm{k}'}{N}}+e^{i2\pi\frac{\mathrm{k}'}{N}})\right]e^{-i2\pi\frac{\mathrm{k}}{N}n} \\
&= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\left[\frac{1}{\sqrt{N}}\sum_{\mathrm{k}'=0}^{N-1}\cos\left(\frac{2\pi\mathrm{k}'}{N}\right)\hat{X}_{\mathrm{k}'}e^{i2\pi\frac{\mathrm{k}'}{N}n}\right]e^{-i2\pi\frac{\mathrm{k}}{N}n} = \frac{1}{N}\sum_{\mathrm{k}'=0}^{N-1}\cos\left(\frac{2\pi\mathrm{k}'}{N}\right)\hat{X}_{\mathrm{k}'}\left[\sum_{n=0}^{N-1}e^{-i2\pi(\frac{\mathrm{k}-\mathrm{k}'}{N})n}\right] \\
&= \frac{1}{N}\sum_{\mathrm{k}'=0}^{N-1}\cos\left(\frac{2\pi\mathrm{k}'}{N}\right)\hat{X}_{\mathrm{k}'}N\delta_{\mathrm{k}\mathrm{k}'} = \cos\left(\frac{2\pi\mathrm{k}}{N}\right)\hat{X}_{\mathrm{k}}.
\end{aligned} \tag{S17}$$

Substituting this result into the Parseval's theorem, we get $\sum_{n=0}^{N-1}X_n\bar{X}_n^* = \sum_{k=0}^{N-1}\hat{X}_{\mathrm{k}}\hat{\bar{X}}_{\mathrm{k}}^* = \sum_{k=0}^{N-1}\cos\left(\frac{2\pi\mathrm{k}}{N}\right)|\hat{X}_{\mathrm{k}}|^2$, where $*$ denotes complex conjugate. So the covariance between X_n and $\bar{X}_n$, $\mathrm{Cov}(X_n,\bar{X}_n)$ is as follows:

$$\begin{aligned}
\mathrm{Cov}(X_n,\bar{X}_n) &= \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[X_n\bar{X}_n^*] - \left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[X_n]\right)\left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[\bar{X}_n]\right) \\
&= \frac{1}{N}\sum_{k=0}^{N-1}\cos\left(\frac{2\pi\mathrm{k}}{N}\right)\mathbb{E}[|\hat{X}_{\mathrm{k}}|^2] - \left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[X_n]\right)\left(\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[\bar{X}_n]\right).
\end{aligned} \tag{S18}$$

By repeating the same procedure for the derivation of $\mathrm{Var}(X_n)$, we get

$$\mathrm{Cov}(X_n,\bar{X}_n) = \frac{\beta^2}{2\pi J}\int_0^\pi \cos u\left[\frac{1-e^{-2Jt(1-\cos u)}}{1-\cos u}\right]du, \tag{S19}$$

and its time derivative is:

$$\frac{d}{dt}\mathrm{Cov}(X_n,\bar{X}_n) = \beta^2 e^{-2Jt}I_1(2Jt). \tag{S20}$$

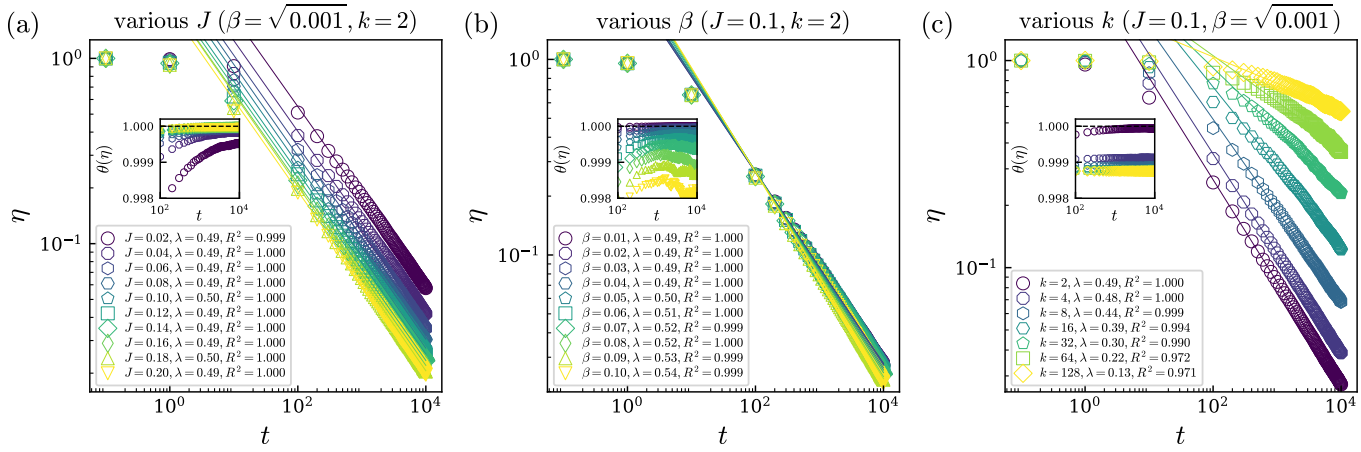

FIG. S2. The field exponent $\eta(t)$ for a variety of the condition (J, β, k). (a) The case with various J and sufficiently small $\beta = \sqrt{0.001}$. (b) The case various β. (c) The case with various k, where k is the number of neighbors in a regular network of 1D ring topology. Open symbols ($\circ$) represent numerical simulation results, and solid lines represent the least square linear regression of data samples for $t \geq 10^2$. R^2 represents the coefficient of determination for regression lines. Insets show $\theta(\eta)$ for $t \geq 10^2$. For all cases, $N = 10^4$, all data are averaged over 128 ensembles.

The indefinite integral is:

$$\begin{aligned}\int d[\mathrm{Cov}(X_n, \bar{X}_n)] &= \beta^2 \int e^{-2Jt} I_1(2Jt) dt = \beta^2 \int e^{-2Jt}(1/2J)\left[\frac{d}{dt} I_0(2Jt)\right] dt \\ &= \frac{\beta^2}{2J}\left[e^{-2Jt} I_0(2Jt) + 2J \int e^{-2Jt} I_0(2Jt) dt\right] = \frac{\beta^2}{2J} e^{-2Jt} I_0(2Jt) + \mathrm{Var}(X_n) + \mathrm{C}. \end{aligned} \tag{S21}$$

Since the initial condition $\mathrm{Cov}(X_n(0), \bar{X}_n(0)) = 0$ and $\lim_{t\to 0} e^{-2Jt} I_0(2Jt) = 1$ give $\mathrm{C} = -\frac{\beta^2}{2J}$. Therefore, the covariance becomes

$$\mathrm{Cov}(X_n, \bar{X}_n) = \mathrm{Var}(X_n) + \frac{\beta^2}{2J}\left[e^{-2Jt} I_0(2Jt) - 1\right]. \tag{S22}$$

The correlation between X_n and X_{n+1}, $\mathrm{Corr}(X_n, X_{n+1})$ is as follows:

$$\mathrm{Corr}(X_n, X_{n+1}) \equiv \frac{\mathrm{Cov}(X_n, X_{n+1})}{\sqrt{\mathrm{Var}(X_n)\mathrm{Var}(X_{n+1})}} = \frac{\mathrm{Cov}(X_n, \bar{X}_n)}{\mathrm{Var}(X_n)}. \tag{S23}$$

where $\mathrm{Cov}(X_n, \bar{X}_n) = (1/2)\left[\mathrm{Cov}(X_n, X_{n-1}) + \mathrm{Cov}(X_n, X_{n+1})\right] = \mathrm{Cov}(X_n, X_{n+1})$ and $\mathrm{Var}(X_n) = \mathrm{Var}(X_{n+1})$ due to the translational invariance of the system. For small t, $c \simeq 1 + x$, so that $\mathrm{Var}(X_n) \simeq \mathrm{Var}(c_n)$ and $\mathrm{Cov}(X_n, \bar{X}_n) \simeq \mathrm{Cov}(c, \bar{c}_n)$. We note that the linear increase of variance and correlation for small t was also revealed by Medo [1]. Our theory for the BM model in a 1D ring not only supports the earlier result but also predicts variance, covariance, and correlation for large t.

From the results of $\mathrm{Var}(X_n)$ and $\mathrm{Cov}(X_n, \bar{X}_n)$, we get

$$1 - \eta = \frac{\mathrm{Cov}(X_n, \bar{X}_n)}{\mathrm{Var}(X_n)} = \frac{\mathrm{Var}(X_n) + \frac{\beta^2}{2J}\left[e^{-2Jt} I_0(2Jt) - 1\right]}{\mathrm{Var}(X_n)} = 1 + \frac{e^{-2Jt} I_0(2Jt) - 1}{2Jte^{-2Jt}\left[I_0(2Jt) + I_1(2Jt)\right]}. \tag{S24}$$

As a result, we can calculate the field exponent $\eta(t)$:

$$\eta(t) = \frac{1 - e^{-2Jt} I_0(2Jt)}{2Jte^{-2Jt}\left[I_0(2Jt) + I_1(2Jt)\right]}. \tag{S25}$$

Since $\eta(t) = -\frac{1}{2J}\frac{I_0(2Jt)}{t[I_0(2Jt)+I_1(2Jt)]} + \frac{\beta^2}{2J}\frac{1}{\mathrm{Var}(X_n)}$ and $\lim_{t\to\infty}\frac{I_0(2Jt)}{t[I_0(2Jt)+I_1(2Jt)]} = 0$, for large t, $\eta(t) = \frac{\beta^2}{2J}\frac{2Ja_0}{\beta^2} t^{-1/2} = a_0 t^{-1/2}$ which matches our numerical simulation results for small β condition.

Figure S2 shows the time evolution of η for a variety of the condition (J, β, k). For small β, Eq. (S25) tells us that the power-law decaying exponent $\lambda = 1/2$ is independent of J. Since $a_0(J) = \sqrt{\pi/(4J)}$, an increase in J is only attributed to the effect of parallel shifting η in double-logarithmic plots. However, for large β, $\lambda > 1/2$. For this case, the difference in X between neighbors cannot be approximated to the first order because it is more affected by multiplicative noise in Eq. (S1). Thus, the variance grows faster than the order of $\sqrt{t}$. The number of neighbors k also can change the value of $\eta(t)$. For sufficiently small k, $\eta(t)$ still exhibits a power-law decay with $\lambda < 1/2$, which is no longer valid for large k. This was also reported in study by Ma et al. [2] as follows: If k is large enough, η has a non-zero finite value and $\rho_{\rm eq}(c)$ satisfying the Fokker-Planck equation is the generalized inverse-gamma distribution. As a result, $\eta(t) \sim a_0 t^{-\lambda}$ for sufficiently small β and k. Thus, the given SDE in Eq. (S1) and Eq. (A3) of *Appendix A* in EM of the main text can be approximated by the time-dependent *Ornstein-Uhlenbeck* process through the EFT ansatz as in Eq. (A4) of *Appendix A* in EM of the main text for such conditions. Therefore,

$$\mathrm{Var}(X_n) = \begin{cases} \beta^2 t & \text{for small } t, \\ \beta^2/(2J\eta) = \beta^2 t^{\lambda}/(2Ja_0) & \text{for large } t, \end{cases} \tag{S26}$$

where $1/2 \le \lambda \le 1$ for the 1D case. For small β, we observe that $\lambda = 1/2$ in Eq. (S25), and for large β, the interaction term becomes relatively small, compared to the fluctuation term, which makes the system behave closer to normal diffusion, so that λ does not exceed 1 (the value for the ballistic behavior).

We conclude this subsection (I A) with a discussion on the mean of x. In fact, deterministic drifts to all X_n only give a translation to the probability density function but do not change its shape. Thus, $\mathrm{Var}(X_n)$ is independent of any homogeneous drift term, so that we arbitrarily define $X_n = \ln c_n + \beta^2 t/2$ in Eq. (S4). Since the normalized income is defined as $c \equiv C/\bar{C}$, it satisfies $\langle c \rangle = 1$. In addition, x follows a log-normal distribution, its mean is $\exp(\mu_t + \sigma_t^2/2)$, and the corresponding drift is $\mu_t = -\sigma_t^2/2$ as we mentioned in the main text as well as EM of the main text. Therefore, the drift of x is not derived by the underlying equation in Eq. (S4), but rather by the normalization drift that c is rescaled at every step.

B. Long-range correlation analysis

The covariance $\mathrm{Cov}(X_n, \bar{X}_n)$ in the previous subsection (I A) is the log-income correlation between neighboring nodes. We here go further and consider correlations between nodes located at greater distances. For any node n, let $\bar{X}_{(n,\ell)}$ be the average log-income of nodes of distance ℓ, then $\bar{X}_{(n,\ell)} = (X_{n-\ell} + X_{n+\ell})/2$. Using the same process in I A, we get

$$\frac{d}{dt}\mathrm{Cov}(X_n, \bar{X}_{(n,\ell)}) = \beta^2 e^{-2Jt} I_\ell(2Jt). \tag{S27}$$

The function $e^{-2Jt} I_\ell(2Jt)$ on the right-hand side has a leading order of $(2Jt)^\ell$ for small t. The correlation between nodes separated by a distance ℓ is as follows:

$$\mathrm{Corr}(X_n, X_{n+\ell}) = \frac{\mathrm{Cov}(X_n, X_{n+\ell})}{\sqrt{\mathrm{Var}(X_n)\mathrm{Var}(X_{n+\ell})}} = \frac{\mathrm{Cov}(X_n, \bar{X}_{(n,\ell)})}{\mathrm{Var}(X_n)}. \tag{S28}$$

We note that $\mathrm{Var}(X_n) = \mathrm{Var}(X_{n+\ell})$ by the translational invariance and be careful for $\mathrm{Cov}(X_n, X_{n+\ell}) = \mathrm{Cov}(X_n, \bar{X}_{(n,\ell)})$ but $\mathrm{Corr}(X_n, X_{n+\ell}) \neq \mathrm{Corr}(X_n, \bar{X}_{(n,\ell)})$. For small t, $\mathrm{Cov}(X_n, \bar{X}_{(n,\ell)}) \sim t^{\ell+1}$ and a leading term of $\mathrm{Var}(X_n) \sim t$, so that the correlation becomes

$$\mathrm{Corr}(X_n, X_{n+\ell}) \sim t^\ell, \quad \text{for } t \ll 1. \tag{S29}$$

Since the initial condition $c_n(0) = 1$ at $t = 0$, the first-order approximation is valid for small t. Thus, $\mathrm{Corr}(c_n, c_{n+\ell}) \simeq \mathrm{Corr}(1 + x_n, 1 + \bar{x}_{n+\ell}) = \mathrm{Corr}(x_n, \bar{x}_{n+\ell})$. Therefore, we get

$$\mathrm{Corr}(c_n, c_{n+\ell}) \sim t^\ell, \quad \text{for } t \ll 1. \tag{S30}$$

This result theoretically supports the correlation analysis in the early time, which was reported by Medo [1]. Moreover, we can show the expression of $\mathrm{Corr}(X_n, X_{n+\ell})$ for large t. As a result, Eq. (S4) is rewritten as a multi-dimensional

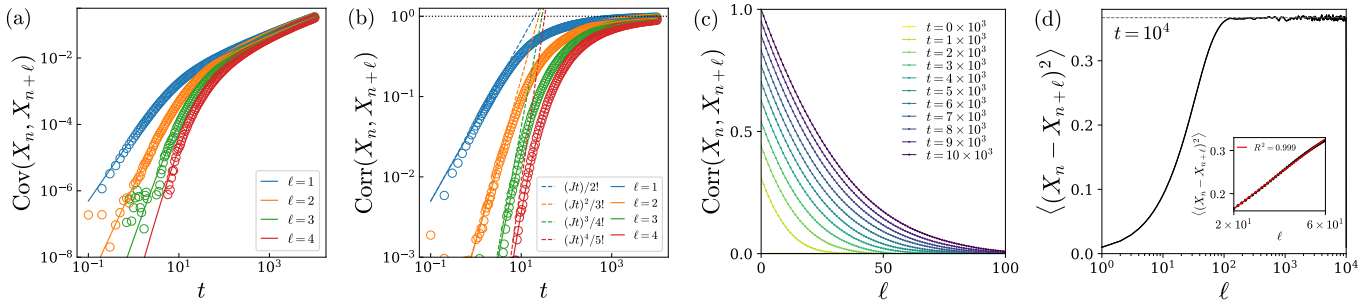

FIG. S3. (a) $\mathrm{Cov}(X_n, X_{n+\ell})$ and (b) $\mathrm{Corr}(X_n, X_{n+\ell})$ against time t for $\ell = \{1, 2, 3, 4\}$. For (a) and (b), open symbols ($\circ$) are numerical simulation results, solid lines are theoretical predictions, and dashed lines are the leading order of the Taylor expansion in Eq. (S35) at $t = 0$. (c) $\mathrm{Corr}(X_n, X_{n+\ell})$ for $\ell \in [0, 100]$ and (d) $\langle (X_n - X_{n+\ell})^2 \rangle$ for $\ell \in [0, 10^4]$. The inset of (d) shows $\langle (X_n - X_{n+\ell})^2 \rangle$ for $\ell \in [20, 60]$. For (c) and (d), solid lines are numerical simulation results. For all cases, $N = 10^4, \alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$, and all data are averaged over 128 ensembles.

OU process:

$$d\mathbf{X}_t = \mathbb{K}\mathbf{X}_t dt + \beta d\mathbf{W}_t, \tag{S31}$$

where $\mathbf{X}_t = (X_0, \ldots, X_{N-1})$, $\mathbf{W}_t = (W_{t,0}, \ldots, W_{t,N-1})$ and $\mathbb{K}$ is a $N \times N$ stability matrix:

$$\mathbb{K} = J \begin{bmatrix} -1 & 1/2 & 0 & \cdots & 1/2 \\ 1/2 & -1 & 1/2 & \cdots & 0 \\ 0 & 1/2 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/2 & 0 & 0 & \cdots & -1 \end{bmatrix}, \tag{S32}$$

The stability of the system is determined by the eigenvalues of $\mathbb{K}$. There are three cases: (1) Stable case, $\mathrm{Re}(\lambda_n) < 0$ for all n; (2) Marginally stable case, $\mathrm{Re}(\lambda_n) \leq 0$ and at least one $\lambda_n = 0$; (3) Unstable case, $\mathrm{Re}(\lambda_n) > 0$ for at least one n. Since $\mathbb{K}$ is circulant matrix, eigenvalues are as follows:

$$\lambda_n = -J\left(1 - \cos\frac{2\pi n}{N}\right) \leq 0 \quad (n = 0, 1, \ldots, N-1). \tag{S33}$$

Thus, $\mathbb{K}$ is marginally stable [3]. For this case, the system shows a strongly auto-correlated behavior for long-time scales. The ℓ-ranged covariance can be obtained by a definite integral of Eq. (S27). Denoting $Q_\ell = \beta^2 \int_0^t e^{-2Js} I_\ell(2Js) ds$, a covariance matrix $\mathbf{\Sigma}_t$ is as follows:

$$\mathbf{\Sigma}_t = \beta^2 \begin{bmatrix} Q_0 & Q_1 & \cdots & Q_{N/2-1} & Q_{N/2} & Q_{N/2+1} & \cdots & Q_1 \\ Q_1 & Q_0 & \cdots & Q_{N/2-2} & Q_{N/2-1} & Q_{N/2} & \cdots & Q_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ Q_1 & Q_2 & \cdots & Q_{N/2} & Q_{N/2-1} & Q_{N/2-2} & \cdots & Q_0 \end{bmatrix}, \tag{S34}$$

where the expression for N is even. For the same time t, we check that $\mathrm{Cov}(X_n, X_{n+\ell})$ decreases against ℓ. The analytical expressions of the covariance for $\ell = \{1, 2, 3, 4\}$ are as follows:

$$\mathrm{Cov}(X_n, X_{n+\ell}) = \begin{cases} \frac{\beta^2}{2J}\left[-1 + (1+2Jt)e^{-2Jt} I_0(2Jt) + 2Jt e^{-2Jt} I_1(2Jt)\right] & (\ell = 1) \\ \frac{\beta^2}{2J}\left[-2 + (2+2Jt)e^{-2Jt}[I_0(2Jt) + I_1(2Jt)]\right] & (\ell = 2) \\ \frac{\beta^2}{2J}\frac{1}{2Jt}\left[-3(2Jt) + 2Jt(5+2Jt)e^{-2Jt} I_0(2Jt) + [-4 + 2Jt(4+2Jt)]e^{-2Jt} I_1(2Jt)\right] & (\ell = 3) \\ \frac{\beta^2}{2J}\frac{1}{(2Jt)^2}\left[-4(2Jt)^2 + 16e^{-2Jt} I_1(2Jt) + 2Jt[-8 + 2Jt(8+2Jt)]e^{-2Jt}[I_0(2Jt) + I_1(2Jt)]\right] & (\ell = 4) \end{cases} \tag{S35}$$

$\mathrm{Corr}(X_n, X_{n+\ell})$ in Eq. (S35) has a leading order as $\sim (Jt)^\ell/(\ell+1)!$ in the early time. Moreover, Eq. (S35) matches for large t, so that we can investigate the case for larger distances ℓ. The long-range spatial correlation could be shown by the spatial variogram $\mathcal{V}(\ell) \equiv \langle (X_n - X_{n+\ell})^2 \rangle$. Since $\mathcal{V}(\ell)$ increases, in proportional to $\sim \log \ell$, until it saturates for

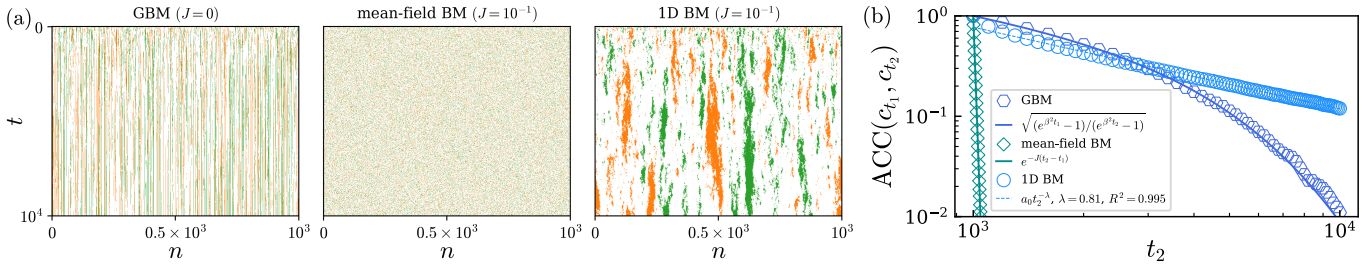

FIG. S4. (a) Spatiotemporal patterns of top-rich/bottom-poor 10% (orange/green) class in homogeneous growth-rate models: the geometric Brownian motion (GBM) in the left panel, the mean-field BM model in the middle panel, and the 1D BM model in the right panel. (b) $\mathrm{ACC}(c_{t_1}, c_{t_2})$ for the three models. Open symbols ($\circ$) represent numerical simulation results, solid lines represent theoretical predictions, and the dashed line is the regression line. Here $t_1 = 10^3$. For all cases, $\alpha = 10^{-2}$ and $\beta^2 = 10^{-3}$. ACCs are obtained by $N = 10^4$ with 128 ensembles.

large ℓ [see the inset of Fig. S3 (d)], the 1D BM model results in the long-range spatial correlation with large clusters in poor and rich classes, compared to other models [see Fig. S4 (a)].

C. Autocorrelation coefficient for various network topologies

We investigate the persistence of income in the BM model by the autocorrelation coefficient [4] (ACC), and compare it with other two versions: (1) the geometric Brownian motion (GBM, non-interactive case) and (2) the mean-field BM model (fully-connected case). The ACC is defined as:

$$\mathrm{ACC}(c_{t_1}, c_{t_2}) \equiv \frac{\mathrm{Cov}(c_{t_1}, c_{t_2})}{\sqrt{\mathrm{Var}(c_{t_1})\mathrm{Var}(c_{t_2})}}. \tag{S36}$$

In the GBM, by using

$$C_t = C(0)\exp\left[\left(\alpha - \beta^2/2\right)t + \beta W_t\right] \quad \text{and} \quad \mathbb{E}[e^{X+Y}] = \exp\left[\mu_X + \mu_Y + (\mu_X^2 + \mu_Y^2)/2 + \mathrm{Cov}(X,Y)\right]$$

for dependent Gaussian random variables X and Y, we get the following statistics: (1) $\mathrm{Var}(c_t) = e^{2\alpha t}(e^{\beta^2 t} - 1)$, and (2) $\mathrm{Cov}(c_t, c_s) = e^{\alpha(t+s)}\left[e^{\beta^2 \min(t,s)} - 1\right]$. Thus, $\mathrm{ACC}(c_t, c_s) = \sqrt{(e^{\beta^2 s} - 1)/(e^{\beta^2 t} - 1)}$ where $t > s$ and $\mathrm{ACC}(c_t, c_s) \approx e^{-\beta^2|t-s|/2}$ for large (t, s). For the mean-field BM model, ACC is known as $e^{-J\Delta t}$ for the condition of $2J/\beta^2 > 1$ [5].In contrast to these cases, which show exponential decay of ACC, for the 1D ring topology, ACC exhibits a power-law decay [see Fig. S4 (b)] with long-term persistence of income.

II. STATISTICAL PROPERTIES OF BINARY MIXTURE IN 1D RING

In this section, we provide all the details for the statistical properties of binary growth rates, $\alpha_\pm = \alpha \pm \Delta\alpha$, in a 1D ring. Here $\alpha > \Delta\alpha > 0$.

A. Assortativity $\mathcal{A}$

We define the growth rate assortativity as $\mathcal{A} \equiv \mathrm{Cov}(\alpha, \alpha')/\sqrt{\mathrm{Var}(\alpha)\mathrm{Var}(\alpha')}$, where α and α' are the growth rates of neighboring nodes and $\alpha \in \{\alpha - \Delta\alpha, \alpha + \Delta\alpha\}$. If we let $N_\pm = N/2$ be the number of nodes with a lower (higher) growth rate, we simply find that $\mathrm{Var}(\alpha) = \mathrm{Var}(\alpha') = \frac{1}{N}\left[\frac{N}{2}(\alpha - \Delta\alpha)^2 + \frac{N}{2}(\alpha + \Delta\alpha)^2\right] - \alpha^2 = \Delta\alpha^2$. For a 1D ring topology, the number of links is the same as the number of nodes N. The number of nodes can be decomposed as $N = N_{ll} + N_{hh} + N_{lh}$, where N_{ll} (N_{hh}) is the number of links between both lower (higher) α nodes and N_{lh} is the

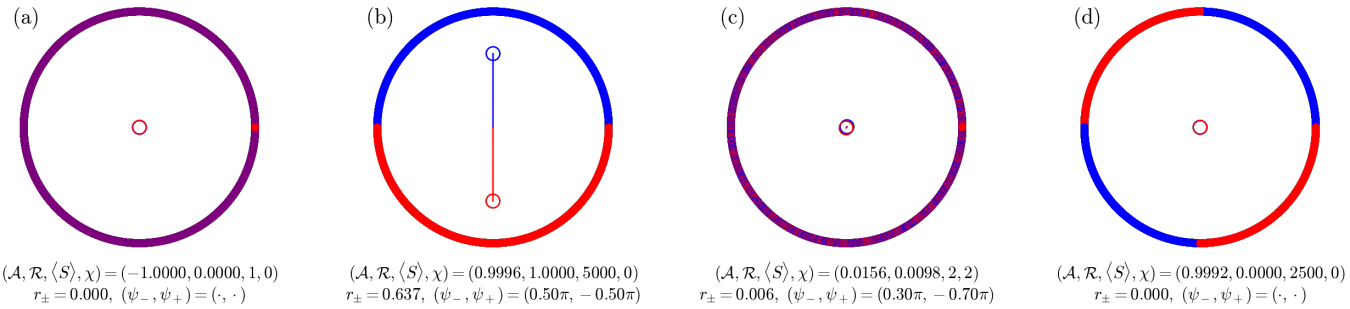

FIG. S5. Statistical properties $(\mathcal{A},\mathcal{R},\langle S\rangle,\chi)$ and order parameters (r,ψ) for a binary mixture in ring topologies: (a) Perfectly disassortative configuration $\mathcal{A}_{\min}$. (b) Perfectly assortative configuration $\mathcal{A}_{\max}$. (c) Randomly permuted configuration. (d) Regularly clustered configuration of $\langle S\rangle=2.5\times10^3$. For all cases, $N=10^4$ and $N_\pm=5\times10^3$.

number of links between the lower and higher α nodes. For this case,

$$\begin{aligned}\mathrm{Cov}(\alpha,\alpha') &= \frac{1}{N}\left[N_{ll}(\alpha-\Delta\alpha)^2+N_{hh}(\alpha+\Delta\alpha)^2+N_{lh}(\alpha-\Delta\alpha)(\alpha+\Delta\alpha)\right]-\alpha^2\\ &= \frac{N_{ll}+N_{hh}-N_{lh}}{N}\Delta\alpha^2+(N_{hh}-N_{ll})(2\alpha\Delta\alpha)=\frac{N^{(1)}-N^{(2)}}{N}\Delta\alpha^2. \end{aligned}\tag{S37}$$

To reflect the changes in the link configuration for each pair of node swapping, we must consider two neighboring values α of the selected nodes. Each node can be categorized by three consecutive network motifs. For this case, there are six motifs and the possible cases of center node swapping between motifs are $\binom{6}{2}$. For all possible cases, $\Delta N_{ll}=\Delta N_{hh}$, so $N_{ll}-N_{hh}$ is conserved under any pair node swapping. Therefore, for the case of $N_-=N_+$, $N_{ll}=N_{hh}$. If we let homogeneous and heterogeneous link densities be $\rho^{(1)}=(N_{ll}+N_{hh})/N$ and $\rho^{(2)}=N_{lh}/N$, then $\mathcal{A}=\rho^{(1)}-\rho^{(2)}$ for the case of $N_-=N_+$.

B. Kuramoto oscillator order parameters (r,ψ): Concentration $\mathcal{R}$

We consider the case of $N_-=N_+$, where the number of each binary element is the same. For a perfectly assortative configuration $(\mathcal{A}_{\max})$ with large N, we can calculate (r_-,ψ_-) and (r_+,ψ_+) as follows:

$$r_-e^{i\psi_-}=\frac{1}{N/2}\sum_{j=1}^{N/2}e^{i(2\pi j/N)}=\frac{1}{2\pi}\frac{2\pi}{N/2}\sum_{j=1}^{N/2}e^{i(2\pi j/N)};\quad \frac{1}{\pi}\int_0^{\pi}e^{iu}du=\frac{2}{\pi}i,\tag{S38}$$

$$r_+e^{i\psi_+}=\frac{1}{N/2}\sum_{j=N/2+1}^{N}e^{i(2\pi j/N)}=\frac{1}{2\pi}\frac{2\pi}{N/2}\sum_{j=N/2+1}^{N}e^{i(2\pi j/N)};\quad \frac{1}{\pi}\int_{\pi}^{2\pi}e^{iu}du=-\frac{2}{\pi}i.\tag{S39}$$

Based on these results, we get $(r_-,\psi_-)=\left(\frac{2}{\pi},\frac{\pi}{2}\right)$ and $(r_+,\psi_+)=\left(\frac{2}{\pi},\frac{3\pi}{2}\right)$, where $2/\pi$ is the maximum value of $r_\pm$. For a perfectly assortative configuration, we show that the following relations are satisfied:

$$r_-=r_+,\quad \Delta\psi=|\psi_--\psi_+|=\pi.\tag{S40}$$

The whole configuration can be generated by pair node swapping that starts from a perfectly assortative configuration. Thus, if the relations of Eq. (S40) are robust under arbitrary pair node swappings, we can address that they are universal properties for the HBM model with a binary mixture of growth rates in the 1D ring.

For an initial binary configuration, its order parameters $\vec{r}_-,\vec{r}_+$ are represented by the real vector form. We assume that the initial configuration satisfies $\vec{r}_-=-\vec{r}_+$. When arbitrary nodes (j,k) are swapped, there are two cases: (1) (j,k) are in the same binary element group and (2) (j,k) are in different binary element groups. For the first case, order parameters do not change. For the second case, the order parameters of the swapped configuration are:

$$\vec{r}'_-=\vec{r}_--\vec{a}_j+\vec{a}_k=\vec{r}_-+\vec{A},\tag{S41}$$

$$\vec{r}'_+=\vec{r}_++\vec{a}_j-\vec{a}_k=\vec{r}_+-\vec{A},\tag{S42}$$

where $\vec{a}_j = \left(\cos\frac{2\pi j}{N}, \sin\frac{2\pi j}{N}\right)$ and $\vec{A} = \vec{a}_k - \vec{a}_j$. For the case of $\vec{r}_- = -\vec{r}_+$, $\vec{r}'_+ = \vec{r}_+ - \vec{A} = -(\vec{r}_- + \vec{A}) = -\vec{r}'_-$. Thus, $\vec{r}'_+$ has the same length and the opposite direction of $\vec{r}'_-$. Therefore, Eq. (S40) is satisfied under arbitrary pair node swapping, except the cases of $\vec{r}'_- = 0$ or $\vec{r}'_+ = 0$. *Path 1* and *Path 2* in the main text [see Fig. 1 (d)] satisfies statistical properties, only except for a perfectly disassortative configuration $(\mathcal{A}_{\min}, \mathcal{R}_{\min})$. There are several cases that satisfy $r = 0$, excluding a perfectly disassortative configuration. If the same binary elements are allocated in exactly opposite directions for all locations, $r = 0$. For this case, the angular argument ψ could not be defined.

C. Cluster-size distribution $P(S)$, $\langle S\rangle$, and χ

The size S distribution of the clusters, $P(S)$, can be characterized by the average $\langle S\rangle$ and the variance χ. For the case of $N_- = N_+$ as $N \to \infty$, a binary mixture on a 1D ring is just a random sequence of binary elements with the same probability $p = 1/2$. If a cluster is defined as a consecutive sequence of the same binary element, the probability mass function of S is as follows:

$$P(S) = \left(\frac{1}{2}\right)^{S+1}, \tag{S43}$$

so that

$$\langle S\rangle = \mathbb{E}[S] = \sum_{S=1}^{\infty} S \cdot P(S) = 2 \quad \text{and} \quad \chi = \mathbb{E}[S^2] - \mathbb{E}[S]^2 = 2$$

for the random configuration as $N \to \infty$. In particular, if all cluster sizes are exactly the same, this configuration satisfies $\mathcal{R} = 0$ with $\chi = 0$ and $N/(2\langle S\rangle) \in \mathbb{N}$. For regularly clustered configurations, $\langle S\rangle$ affects the spatiotemporal patterns of income dynamics.

III. HETEROGENEOUS BM MODEL IN 1D RING

In this section, we provide all the details (analytical derivations and numerical confirmations) for the heterogeneous BM (HBM) model in a 1D ring, which is compared to our findings in the main text.

A. Normalized income distribution $\rho(x)$: Perfectly disassortative *versus* Perfectly assortative

Let $\rho(x), \rho_{\alpha_-}(x)$, and $\rho_{\alpha_+}(x)$ be the probability density functions of x for the entire, the lower and higher α groups, respectively. For a perfectly disassortative growth rate configuration with $\mathcal{A}_{\min}$, $\rho_{\alpha_-}(x)$ and $\rho_{\alpha_+}(x)$ overlap almost perfectly each other, and the overall distribution is almost the same as that in the BM model. Therefore, $\rho(c, t; \mathcal{A}_{\min}) \sim \text{Lognormal}(\mu_t, \sigma_t^2)$. For a perfectly assortative growth rate configuration with $\mathcal{A}_{\max}$, both of $\rho_{\alpha_\pm}(x)$ do not overlap almost perfectly each other. For this case, $\rho(x, t)$ can be roughly divided into three regions: (1) *Head* – the first peak; (2) *Body* – the middle part between the first and second peaks; (3) *Tail* – the second peak. In Fig. S6 (a)-(c), we observe more precisely how many samples exist between the two peaks. Surprisingly, we find that

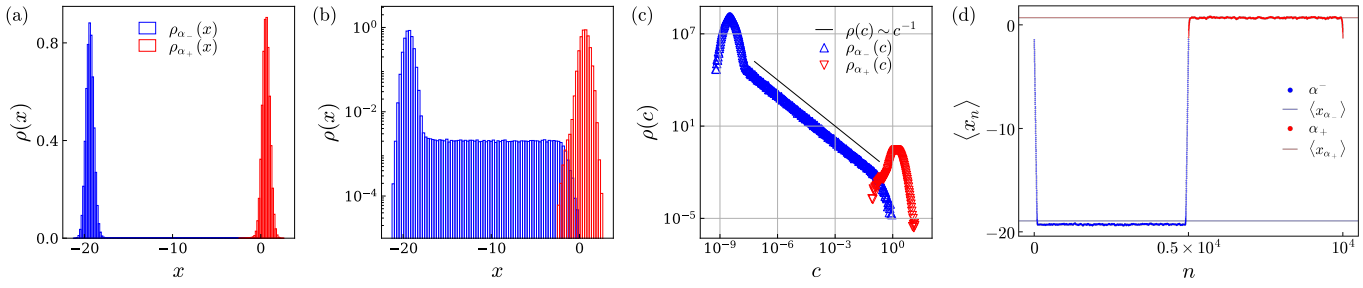

FIG. S6. Log-income ($x = \ln c$) distribution for the case of $\mathcal{A}_{\max}$: (a) $\rho_{\alpha_\pm}(x)$ and (b) the semi-logarithmic scaled plot of $\rho_{\alpha_\pm}(x)$. (c) Double-logarithmic scaled plot of $\rho(c)$, where the black solid line represents the guided line of $\rho(c) \sim c^{-1}$. (d) The positional mean of x, $\langle x_n\rangle$. For all cases, $N = 10^4, \alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}, t = 10^4$ and all data are averaged over 128 ensembles.

the samples of x between the two peaks follow a uniform distribution, where most of them belong to the α_- group. It implies that for the case of $\mathcal{A}_{\max}$, $\rho_{\alpha_-}(x)$ and $\rho_{\alpha_+}(x)$ are asymmetric over x, whereas for the case of $\mathcal{A}_{\min}$, they are symmetric. In Fig. S6 (c), the double logarithmic scaled plots show that the body region follows a power-law of $\rho(c) \sim Ac^{-1}$, where A is a constant. This is a natural consequence of the distribution transformation formula: $\rho(c) = \rho(x)|dx/dc| = A|d(\ln c)/dc| = Ac^{-1}$. Our hypothesis in the regional separation of the distribution function is that the field exponent η depends on the positional index of a node n.

Equation (A6) in *Appendix A* of EM, is obtained by the first-order approximation with the assumption that ηx is sufficiently small. However, if η is extremely small, the interaction term $J\left[\theta(\eta)c^{1-n} - c\right]dt$ in Eq. (A5) can be neglected due to $c^{1-\eta} \to c$ and $\theta(\eta) \to 1$. Therefore, Eq. (A5) becomes separated into two cases as follows:

$$\begin{aligned} & dc_n = J[\theta(\eta)c_n^{1-\eta} - c_n]dt + \beta c_n dW_{t,n}, \\ \longrightarrow & \begin{cases} dx_n &= J\eta_{t,n}[\mu_{t,n} - x_n]dt + \beta dW_{t,n} \\ dc_n &= \beta c_n dW_{t,n} \end{cases} \end{aligned} \tag{S44}$$

If our hypothesis is correct, the governing equation for $\rho(c,t)$ should depend on the node's position n for the given network. In Fig. S6 (d), we plot $\langle x_n \rangle$ against n, which confirms our hypothesis. In the near body region between two different α clusters, $\langle x_n \rangle$ is almost equidistantly spaced and proportional to n. Since $\langle x_n \rangle$ is obtained by the ensemble average at position n, there is the same number of node samples for each value. Therefore, x drawn from these samples of nodes becomes uniformly distributed in $\rho(x)$. Since the number of samples is the same at each point, $\rho^{(\mathrm{b})}(x,t) = \text{const}$. Therefore, $\frac{\partial}{\partial t}\rho^{(\mathrm{b})}(c,t) = 0$.

The Fokker-Planck equation for the second case of Eq. (S44) is as follows:

$$\frac{\partial}{\partial t}\rho(c,t) = \frac{1}{2}\frac{\partial^2}{\partial c^2}\left[\beta^2 c^2 \rho(c,t)\right]. \tag{S45}$$

For the body region, $\frac{\partial}{\partial t}\rho^{(\mathrm{b})}(c,t) = 0$ and its solution is $\rho^{(\mathrm{b})}(c,t) = Ac^{-1} + Bc^{-2}$, where A and B are constants. By numerical simulations, we observe $\rho^{(\mathrm{b})}(c,t) \sim Ac^{-1}$, which supports $B = 0$. Thus, the governing equation for c_n in this region is equal to the second case of Eq. (S44), and η measured from these samples is extremely small. For this case, we call the set of nodes the body class if c follows power-law as $\rho(c) \sim Ac^{-1}$. Automatically, the regions excluding the body class should be the head class, represented by small $\langle x_n \rangle$, and the tail class, represented by large $\langle x_n \rangle$, respectively [see Fig. S6 (d)]. Head and tail classes exhibit power-law decays, which are the same as those in the BM model [see Fig. S7 (e) for the field exponent $\eta(t)$ in each region], while η for the body class decreases much faster. At $t = 10^4$, the order of η is smaller than 10^{-4}. As a result, the interaction term in Eq. (S44) can be neglected. The dynamics of head and tail classes is governed by the first case of Eq. (S44), described by the log-normal distribution, while the dynamics of the body class is governed by the second case of Eq. (S44), which follows a power-law distribution [see Fig. S6 (c)]. The remainder is the drift $\mu_{t,n}$ for head and tail classes. To figure it out, we consider the null model, where the two α groups are completely separated by two 1D rings. This implies that there is no body region. For this case, each group can be treated as an independent BM model, and the underlying equation is the same as Eq. (1) in the main text. The average income C of them becomes $\langle C_{\alpha_\pm}(t) \rangle = C(0)e^{\alpha_\pm t}$. Then, the corresponding average normalized income c for each group is as follows:

$$\langle c_{t,\alpha_\pm} \rangle_{(\text{null})} = \frac{C(0)e^{(\alpha \pm \Delta\alpha)t}}{[C(0)e^{(\alpha-\Delta\alpha)t} + C(0)e^{(\alpha+\Delta\alpha)t}]/2} = \frac{2}{1 + e^{\mp 2\Delta\alpha t}}, \tag{S46}$$

Since each group follows a log-normal distribution that has the variance $\sigma_t^2 = \beta^2 t^\lambda/(2Ja_0)$ for large t, similar to the homogeneous BM model, the normalization conditions are $\langle c_{t,\alpha_\pm} \rangle_{(\text{null})} = \exp\left(\mu_t^\pm + \sigma_t^2/2\right)$. Therefore,

$$\mu_t^\pm \equiv \langle x_{t,\alpha_\pm} \rangle_{(\text{null})} = -\sigma_t^2/2 + \ln\left(\frac{2}{1 + e^{\mp 2\Delta\alpha t}}\right), \quad \Delta\mu \equiv |\mu_t^+ - \mu_t^-| \approx \begin{cases} 0 & \text{for small } t, \\ 2\Delta\alpha t & \text{for large } t. \end{cases} \tag{S47}$$

As a result, the interval between two Gaussian peaks is linearly proportional to t for large t. The larger the difference of two growth rates ($\Delta\alpha$), the wider the segregation of two income levels ($\Delta\mu$).

We empirically find that $\langle c_t^{(\mathrm{h})} \rangle$, $\langle c_t^{(\mathrm{t})} \rangle$, $\langle x_t^{(\mathrm{h})} \rangle$, and $\langle x_t^{(\mathrm{t})} \rangle$ are the same as those of the null model [see Fig. S8 (a)-(d)]. In addition, the segregation of income levels of tail and head nodes, $\Delta\mu' \equiv |\mu_t^{(\mathrm{t})} - \mu_t^{(\mathrm{h})}|$, is compared to $\Delta\mu' - \Delta\mu \sim -b_0 t$, where $b_0 \ll \Delta\alpha$ [see Fig. S8 (e) and the inset]. As a result, $\mu_t^{(\mathrm{h})} \approx \mu_t^-$ and $\mu_t^{(\mathrm{t})} \approx \mu_t^+$. Most of the nodes in the uniform distribution region belong to the α_- group and only a small fraction belongs to the α_+ group. As the distance between

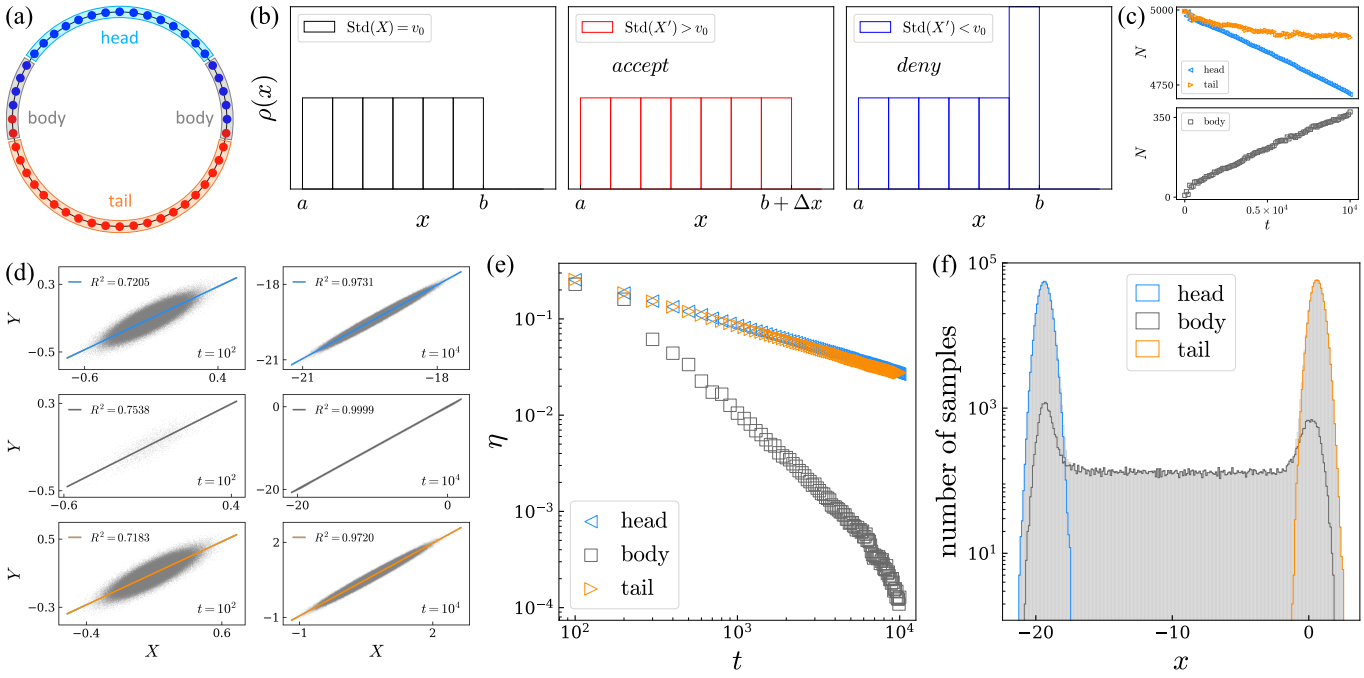

FIG. S7. Positional separation for the case of $\mathcal{A}_{\max}$: (a) Conceptual visualization of three parts for the separation of governing equations in a 1D ring. (b) Standard deviation maximization algorithm. Detecting the body region, we start from an active node (the node that has a heterogeneous link) and select a set of nodes within the range $[d_1, d_2]$ from the active node. The standard deviation of X, $\mathrm{Std}(X)$, from theses node samples is maximized in the range $[d_1, d_2]$. (c) The number of nodes belonging to each region, $\mathcal{N}$ against time t for three regions: [top] tail (orange) and head (blue); [bottom] body (black). (d) Correlation between $X = \ln c$ and $Y = \ln \bar{c}_n$ for three regions at $t = 10^2$ and $t = 10^4$: Head (top), body (middle), and tail (bottom). (e) Field exponent $\eta(t)$ for head (blue), body (black), and tail regions (orange) against t. (f) Portion of x samples for head, body, and tail regions. In (c)-(f), $N = 10^4, \alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$, and all data are averaged over 128 ensembles.

two Gaussian peaks increases as time elapses, the uniform distribution area becomes broader because samples that belong to the Gaussian peaks are absorbed into the uniform distribution region over time. At this moment, in the α_- group, more samples are absorbed than in the α_+ group. Therefore, the relative height difference of two Gaussian peaks grows as the uniform distribution region widens over time. This is due to a completely finite N effect. Figure S7 (c) shows the number of samples for each class, which are summarized as follows [see Eq. (7) in the main text]:

$$\rho^{(\mathrm{h})}(c,t) \sim \text{Lognormal}\left(\mu_t^-, \sigma_t^2\right); \ \rho^{(\mathrm{b})}(c,t) \sim Ac^{-1}; \ \rho^{(\mathrm{t})}(c,t) \sim \text{Lognormal}\left(\mu_t^+, \sigma_t^2\right). \tag{S48}$$

Therefore, we conclude that if growth rates are perfectly segregated in a 1D periodic lattice (ring), the log-income distribution is represented by a mixture of log-normal and power-law distributions, and the income level segregation between two growth rate ($\alpha = \alpha_\pm$) groups increases as $\Delta\mu \approx 2\Delta\alpha t$. We note that the difference in η changes the effective governing equation for c_n [see Eq. (S44)], resulting in two different types of distributions.

B. Gini index g

For the case of $\mathcal{A}_{\max}$, the contribution of body samples to the Gini index g becomes negligible since the number of body samples is very small [see Fig. S6 (a)]. If a probability distribution $\rho(c)$ represents the dual log-normal mixture of Lognormal(μ_t^-, σ_t^2) and Lognormal(μ_t^+, σ_t^2) with the fraction of f_1 and f_2, g is as follows:

$$g = \frac{f_1^2 e^{\mu_t^-}\,\mathrm{erf}(\frac{\sigma_t}{2}) + f_2^2 e^{\mu_t^+}\,\mathrm{erf}(\frac{\sigma_t}{2}) + f_1 f_2\left[e^{\mu_t^-}\,\mathrm{erf}\left(\frac{\mu_t^- - \mu_t^+ - \sigma_t^2}{2\sigma_t}\right) + e^{\mu_t^+}\,\mathrm{erf}\left(\frac{\mu_t^+ - \mu_t^- - \sigma_t^2}{2\sigma_t}\right)\right]}{f_1 e^{\mu_t^-} + f_2 e^{\mu_t^+}}. \tag{S49}$$

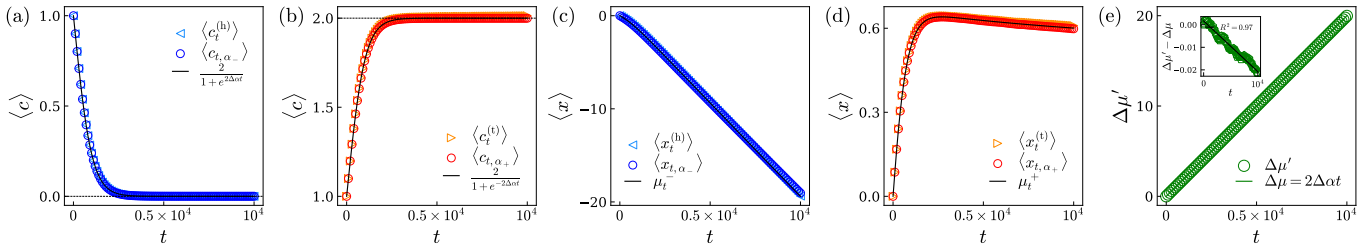

FIG. S8. Ensemble averages and income level segregation for $\mathcal{A}_{\max}$ case. (a), (b) Time evolutions of $\langle c \rangle$ for head-tail classes, and $\alpha_\pm$ groups. (c), (d) Time evolutions of $\langle x \rangle$ for head-tial classes and $\alpha_\pm$ groups. (e) Income level segregation between two Gaussian peaks $\Delta\mu'$ (see Fig. S7-(f)). The inset shows $\Delta\mu' - \Delta\mu$. For all cases, $N = 10^4, \alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$ and results are obtained by 128 ensembles.

For our case, $f_1 = f_2 = 1/2$, $\mu_t^+ > \mu_t^-$, and $(\mu_t^+ - \mu_t^-) \gg \sigma_t^2$ for large t. Thus, $g \approx \frac{1}{2}\left(1 - \frac{2}{1+e^{2\Delta\alpha t}}\right) + \frac{1}{2}\mathrm{erf}\left(\frac{\sigma_t}{2}\right)$ for large t, which is consistent with Eq. (9) in the main text and Eq. (B7) in *Appendix B* [see Fig. 3 (b) in the main text].

C. Income level segregation $\Delta\mu$

For the case of $\mathcal{A}_{\max}$, the income level segregation of the $\alpha_\pm$ groups is $\Delta\mu \approx 2\Delta\alpha t$. However, in general, $\Delta\mu$ depends on the α configurations. For $\mathcal{R} \sim 0$ (*Path 1*), $\rho_{\alpha_\pm}$ overlap almost each other, so that $\Delta\mu \sim 0$ [see Fig. S9 (a) and (c)]. For $\mathcal{R} > 0$ (*Path 2*), the nodes with α_- can be a rich class and vice versa. Thus, $\Delta\mu$ decreases as the overlap of $\rho_{\alpha_\pm}$ increases [see Fig. S9 (b) and (d)]. For those cases, we empirically find that

$$\Delta\mu \sim \begin{cases} 0 & \text{for } \textit{Path 1}\text{: } (\mathcal{A} \le 0, \mathcal{R} \sim 0), \\ \mathcal{A} \times 2\Delta\alpha t & \text{for } \textit{Path 2}\text{: } (\mathcal{A} > 0, \mathcal{R} > 0). \end{cases} \tag{S50}$$

Here $\Delta\mu$ has a very small value for $\mathcal{R} \sim 0$, which is controlled by the assortativity $\mathcal{A}$ for $\mathcal{R} > 0$ [see Fig. S9 (a) and (b)]. We confirm that this linear relationship is almost valid for $t \sim 10^4$. However, for very large t, the distribution becomes irregular, which cannot represent two peaks. This seems to be a finite N effect since $\rho_{\alpha_\pm}(x,t)$ for large t strongly depends on local disorder in the configuration.

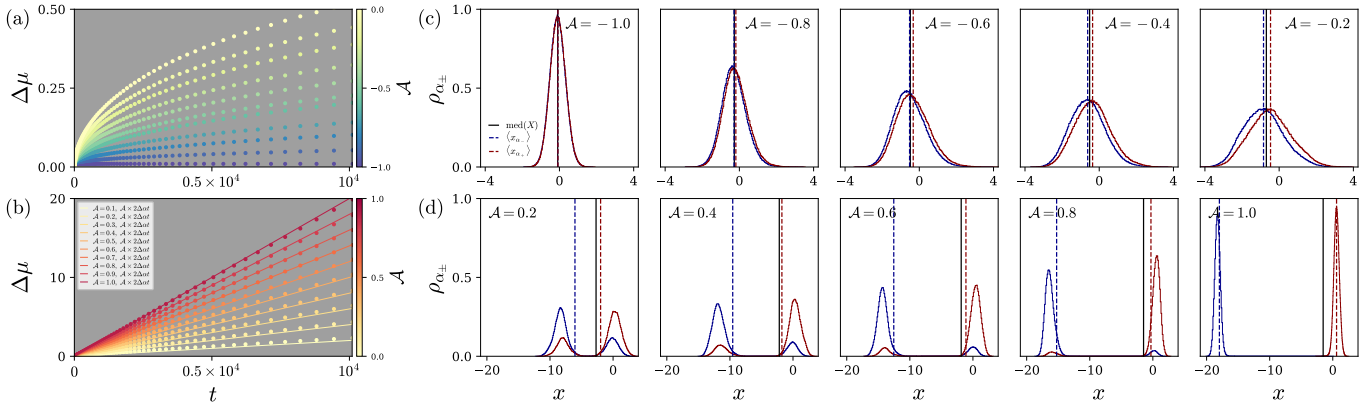

FIG. S9. Income level segregation $\Delta\mu$ and distribution $\rho_{\alpha_\pm}(x)$ for various α configurations chosen from random pair swapping, *Path 1* and *Path 2* [see Fig. 1 in the main text]: For $\Delta\mu$, (a) *Path 1* and (b) *Path 2*. For $\rho_{\alpha_\pm}(x,t)$ at $t = 10^4$, (c) *Path 1* and (d) *Path 2*. For all cases, $\alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$ and all data are averaged over 128 ensembles.

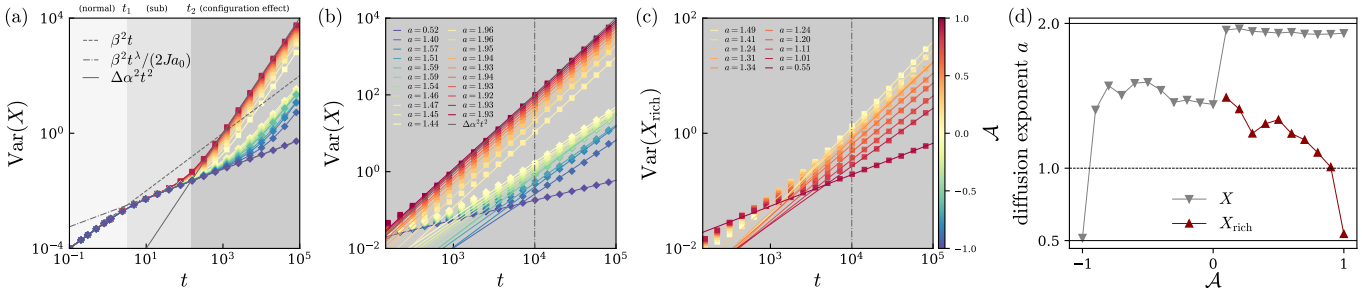

FIG. S10. Variances in the HBM model for various α configurations from random pair swapping, *Path 1* ($\diamond$) and *Path 2* ($\square$) [see Fig. 1 (d) in the main text]. (a) Var(X), (b) Var(X) for $t > t_2$, (c) Var($X_{\rm rich}$) for $t > t_2$ and $\mathcal{A} > 0$, and (d) diffusion exponent a, where $X_{\rm rich}$ is the samples that excess the median of X. Colored solid lines show the linear regression for samples with $t > 10^4$, and the diffusion exponent a is described in the legends. Here $\alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$, and all data are averaged over 128 ensembles.

D. Diffusive nature and ballistic motion: Variance Var(X)

If the variance of X follows a power law as $\mathrm{Var}(X) \sim t^a$, the diffusion exponent a characterizes the anomaly of diffusive behaviors. In the HBM model for a 1D ring topology, the total variance of X passes through three regimes: (1) Normal diffusion, (2) sub-diffusion, and (3) configuration-effect dominant diffusion. $\rho(x, t; \mathcal{A}_{\max})$ can be approximated as a dual Gaussian mixture of $\mathcal{N}(\mu_t^-, \sigma_t^2)$ and $\mathcal{N}(\mu_t^+, \sigma_t^2)$. As a result,

$$\mathrm{Var}(X; \mathcal{A}_{\max}) \approx \sigma_t^2 + \frac{(\mu_t^+ - \mu_t^-)^2}{4} = \begin{cases} \beta^2 t & \text{for small } t, \\ \beta^2 t^\lambda/(2Ja_0) + \Delta\alpha^2 t^2 & \text{for large } t, \end{cases} \tag{S51}$$

where $|\mu_t^+ - \mu_t^-| = \Delta\mu$ in Eq. (S47). Since $0.5 \le \lambda \le 1$, $\mathrm{Var}(X; \mathcal{A}_{\max}) \approx \Delta\alpha^2 t^2$ for very large t.

Figure S10 (a) shows that the variance Var(X) is characterized by triple time scales: $\{\beta^2 t, \beta^2 t^\lambda/(2Ja_0), \Delta\alpha^2 t^2\}$. The corresponding intersections are $t_1 = [2Ja_0]^{1/(\lambda-1)}$ and $t_2 = [2Ja_0\Delta\alpha^2/\beta^2]^{1/(\lambda-2)}$. For $t_1 < t < t_2$, the system deviates from the normal diffusion regime of the GBM and enters the sub-diffusion regimes in the BM model. For $t > t_2$, the system deviates from sub-diffusion and enters the configuration-effect dominant regime of the HBM model. This is the unique feature of the HBM model since the BM model always ends with sub-diffusion. From the analytical form of t_2, we can easily expect that the sub-diffusion regime vanishes if $\Delta\alpha \gg \beta$.

For $\mathcal{R} \sim 0$ (*Path 1*), $\rho_{\alpha_\pm}$ almost overlap each other and Var(X) is enough to capture the diffusive nature of the system [see Fig. S9 (c)]. Not only the case of $\mathcal{A}_{\max}$ but also the other case of $\mathcal{R} > 0$ (*Path 2*) shows that $\mathrm{Var}(X) \sim t^2$ because the distance between probability density peaks increases almost linearly over time t, implying that the ballistic motion of Var(X) is more dominated than diffusion [see Fig. S10 (b) and (d)]. To investigate the diffusive nature of the peak, we must consider only the variance of the single peak. Fortunately, we empirically find that the median of X almost separates the peak of the rich side and the others [see Fig. S9 (d)]. Thus, by investigating Var($X_{\rm rich}$) where $X_{\rm rich} = \{X | X > \mathrm{med}(X)\}$, we can identify the diffusive nature of the peak. We numerically estimate the diffusion exponent a in terms of linear regression for large t, and find that for Var($X_{\rm rich}$), $0.5 < a < 2$ [see Fig. S10 (c) and (d)]. In short, the configurational property $\mathcal{A}$ controls the diffusive nature of the system, and the system lies in the sub-diffusion regime to the super-diffusion regime for large t.

IV. HBM MODEL IN WATTS-STROGATZ NETWORK

In this section, we provide detailed numerical simulations for the HBM model in a Watts-Strogatz (WS) network, where we use the WS network with mean degree $k = 4$, not $k = 2$. We note that the rewiring procedure with $k = 2$ makes several divided components where the income dynamics is not consistent, whereas the case of $k = 4$ does.

Since the WS network is not a ring topology, $\mathcal{R}$ is no longer valid. However, $\mathcal{A}$ is still valid for an arbitrary network. The configuration property $\mathcal{A}$ depends on the initial configuration and the rewiring probability p of the WS network. We use two initial configurations: (1) Alternatively allocated $\mathcal{A}_1$ and (2) fully segregated $\mathcal{A}_2$, which correspond to $\mathcal{A}_{\min}$ and $\mathcal{A}_{\max}$ for the 1D ring case, respectively. $\mathcal{A}_1^{(p)}$ ($\mathcal{A}_2^{(p)}$) as an assortativity after the rewiring process with probability p starts from the initial configuration $\mathcal{A}_1$ ($\mathcal{A}_2$). For the case of $\mathcal{A}_1^{(p)}$, $\mathcal{A}_1^{(0)} = 0$ because the number of homogeneous and heterogeneous links is the same. The additional rewiring with p does not change the number of two

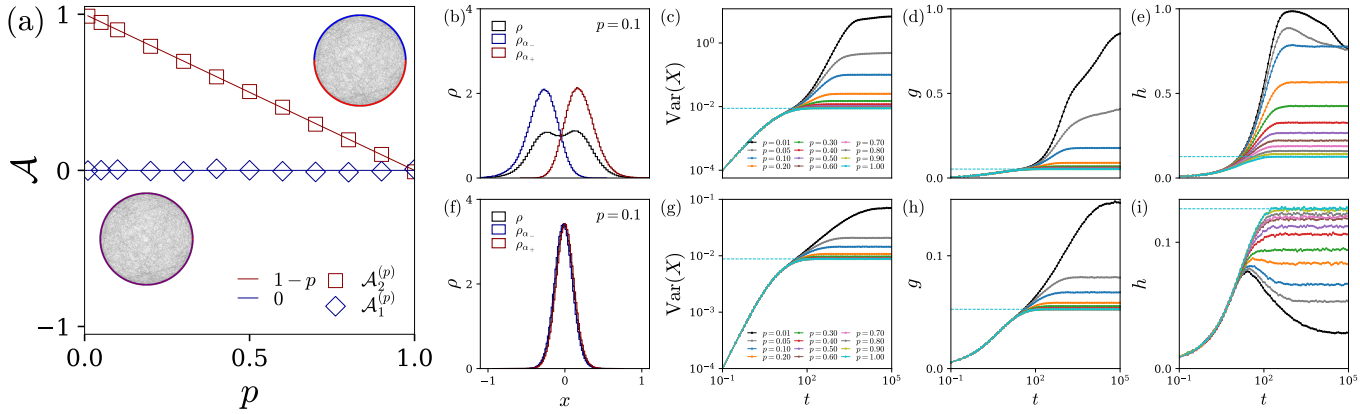

FIG. S11. (a) Configuration properties of $\mathcal{A}_1^{(p)}$ and $\mathcal{A}_2^{(p)}$ with two insets that show two network visualizations for $\mathcal{A}_1^{(p)}$ (upper) and $\mathcal{A}_2^{(p)}$ (lower) with $p = 0.1$, respectively. Log-income distribution $\rho(x)$ at $t = 10^4$ with $p = 0.1$: (b) for $\mathcal{A}_2^{(p)}$ and (c) for $\mathcal{A}_1^{(p)}$. (c)-(e) and (g)-(i) show Var(X), g, and h for $\mathcal{A}_1^{(p)}$ and $\mathcal{A}_2^{(p)}$, respectively. For all cases, $\alpha = 10^{-2}, \Delta\alpha = 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$ and all data are averaged over 128 ensembles.

types of links on average, so that $\mathcal{A}_1^{(p)} \approx 0$ [see Fig. S11 (a)]. For the case of $\mathcal{A}_2^{(p)}$, $\mathcal{A}_2^{(0)} = +1 - 12/(2N)$ because the number of heterogeneous links is 6 and the total number of links is $2N$. For this case, the additional rewiring with p changes assortativity because the number of homogeneous links is proportional to $(p/2)$, finally, the number of two types of links on average is balanced at $p = 1$. Thus, $\mathcal{A}_2^{(p)} \approx 1 - p$ [see Fig. S11 (a)].

In the BM model study by Souma *et al.* [6], the small-world (SW) effect changes $\rho(x)$ from log-normal to power-law, and reduces the Gini index g. In contrast to the 1D case, in the WS network with sufficiently large p, Var(X) saturates for large t, implying that $\rho(x)$ converges to the stationary distribution [see Figs. S1 (a), S10 (a), and S12 (c), (g)]. The HBM model shows a stationary distribution with sufficiently large p and a shift from log-normal to stationary power-law as p increases. However, in the HBM model, both $\mathcal{A}$ and p determine $\rho(x)$, which is differently dependent on $\mathcal{A}$ even for the same p [see Fig. S11 (b) and (f)], and on p even for the same $\mathcal{A}$ [see Fig. S11 (g)-(i)]. In particular, $\mathcal{A}_1^{(p)}$ exhibits a stationary unimodal distribution; however, $\mathcal{A}_2^{(p)}$ with appropriate p exhibits a stationary bimodal distribution, corresponding to the second era of the history of global inequality, reported by Milanovic [7]. For the *"small-networkness"*, we here consider $p \in [10^{-4}, 10^{-1}]$, not $p \in [10^{-1}, 1]$ in the main text.

In the limit of $p \to 0$, the WS network becomes regular, so that the field exponent η is almost the same as that in the 1D BM model [see Fig. S2 (c)]. Thus, the spatiotemporal patterns of income dynamics are also almost the same [see Fig. S12 (a) and Fig. 1 (c) in the main text]. For $10^{-4} < p < 10^{-2}$, $h \sim 1$ represents that $\rho_{\alpha_\pm}$ are almost perfectly decoupled, and $g > 1/2$. It is because the between-inequality in $\alpha_\pm$ groups guarantees half of the Gini index g. In this region, the larger p, the larger g because p makes heterogeneous links change $\rho_{\alpha_\pm}$ more diffusive, so that the within-inequality increases. However, for $p > 10^{-2}$, both the segregation of $\alpha_\pm$ groups and the variance for each peak decrease, so that both between- and within-inequalities decrease, represented by a rapid decrease in g [see Fig. S12 (e)]. For the case of the WS network, in contrast to the the 1D case, $\Delta\mu$ depends not only on $\Delta\alpha$ but also

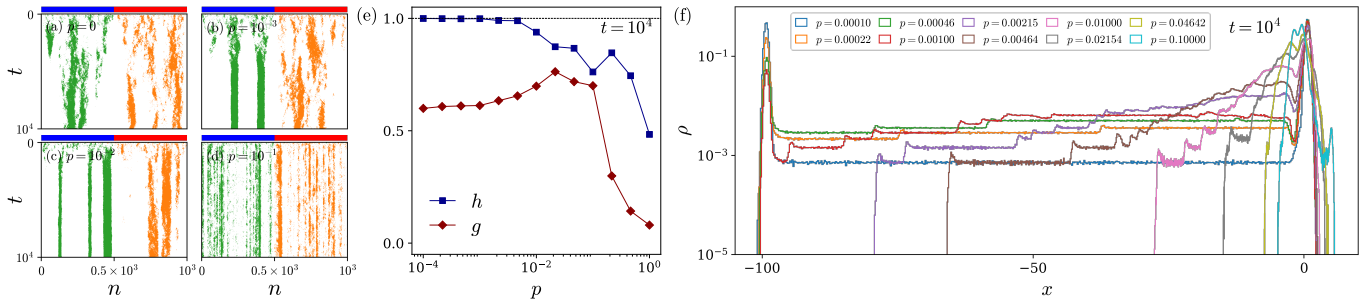

FIG. S12. SW effect on HBM model: (a)-(d) Spatiotemporal patterns of top-rich/bottom-poor 10% (orange/green) class for $\mathcal{A}_2^{(p)}$ for $p = \{0, 10^{-3}, 10^{-2}, 10^{-1}\}$. (e) Gini index g and Hellinger distance h against $p \in [10^{-4}, 10^{-1}]$. (f) $\rho(x)$ against log-income x for $p \in [10^{-4}, 10^{-1}]$. For all cases, $\alpha = 10^{-2}, \Delta\alpha = 5 \times 10^{-3}, \beta^2 = 10^{-3}, J = 10^{-1}$, and all data averaged over 128 ensembles at $t = 10^4$.

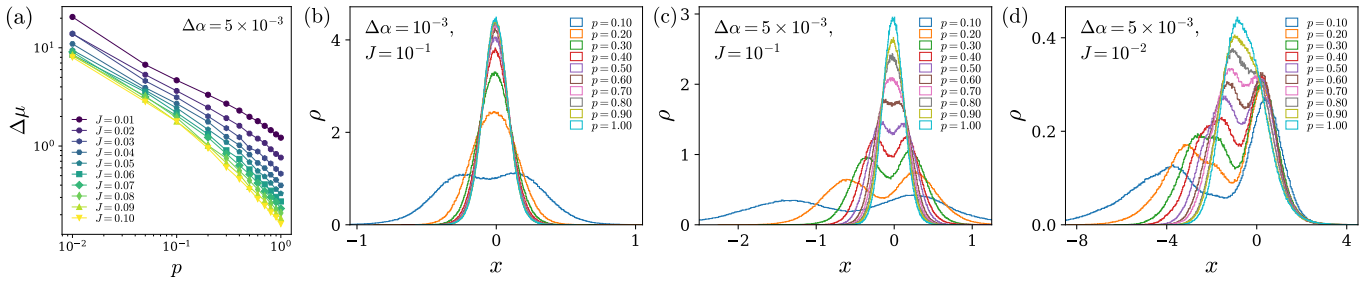

FIG. S13. Parameter dependence on HBM model in WS network: (a) J dependence on $\Delta\mu$ against p for $\Delta\alpha = 5\times10^{-3}$ and $J = \{0.01, 0.02, \cdots, 0.10\}$ from top to bottom. (b)-(d) $(\Delta\alpha, J)$ dependence on log-income distribution $\rho(x)$ against x: (b) for $(\Delta\alpha, J) = (10^{-3}, 10^{-1})$, (c) for $(\Delta\alpha, J) = (5\times10^{-3}, 10^{-1})$, and (d) for $(\Delta\alpha, J) = (5\times10^{-3}, 10^{-2})$. For all cases, $N = 10^4, \alpha = 10^{-2}, \beta^2 = 10^{-3}$ and all data are averaged over 128 ensembles at $t = 10^4$.

on J [see Eq. (S50) and Fig. S13].

As a final remark, we briefly address that the Newman-Watts (NW) model also yields almost similar results as those by the WS network since the stationarity and the shape of distributions mainly depend on the average shortest path length in such a network. Moreover, we confirm that the NW network with the same parameters as before shows a stationary bimodal log-income distribution, where other results are also consistent with the case of the WS network.

V. PHYSICAL INTERPRETATIONS OF OUR STUDY

In this section, we highlight the conceptual advance in our study related to the universality of local interactions.

For the 1D BM model, the continuum limit of the interaction term in Eq. (A1) becomes a spatial diffusion term (or Laplacian). In that case, the equation becomes a 1D stochastic heat equation (SHE) with multiplicative noise. It is known that the Cole-Hopf transformation, which takes the logarithm, gives the 1D Kardar-Parisi-Zhang (KPZ) equation. Thus, our log-income ($x \equiv \ln c$) of each region can be considered as a KPZ height field. Also, we have shown that if fluctuation β is small, multiplicative noise can be transformed by additive noise, meaning that log-income can be considered as a height field of the Edwards-Wilkinson (EW) equation.

For the HBM model, we introduced heterogeneous growth rates $\alpha_n \in \{\alpha - \Delta\alpha, \alpha + \Delta\alpha\}$ as a quenched disorder, and we still consider space-time noise represented by the Wiener process of each region. Therefore, our model can be mapped into a quenched KPZ equation with space-time noise.

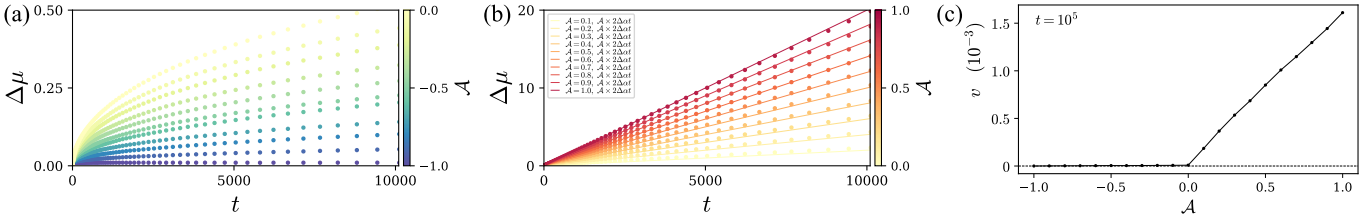

FIG. S14. (a), (b) Income level segregation $\Delta\mu$ and (c) average segregation velocity v.

In contrast to quenched 1D-KPZ and 1D-EW equations, the degree of driving force F is not an control parameter since growth rates α_n are given and fixed. The important thing is that in our case, configurational properties play as control parameters for a Pinning-Depinning (PD)-like transition. As we have shown in Fig. S9, income level segregation $\Delta\mu$ sub-linearly increases for *Path 1* configurations but linearly increases for *Path 2* configurations. If we define average segregation velocity $v \equiv \Delta\mu/t$, v shows continuous phase transition over $\mathcal{A}$ [see Fig. S14-(c)].

$$v \equiv \Delta\mu/t \sim \begin{cases} 0 & \text{for } \textit{Path 1}\text{: } (\mathcal{A} \le 0, \mathcal{R} \sim 0), \\ \mathcal{A} \times 2\Delta\alpha & \text{for } \textit{Path 2}\text{: } (\mathcal{A} > 0, \mathcal{R} > 0). \end{cases} \tag{S52}$$

It shows that height segregation of the surface emerges from the configurational properties of growth rates.

Since we simultaneously consider quenched disorder and space-time noise, there are different growth patterns from the 1D-KPZ or 1D-EW equation. As we have shown in Fig. S10, various configurational properties give different

diffusion exponents (twice the growth exponent in the surface growth model). It means that the coexisting interplay of quenched disorder and space-time noise makes an unseen growth exponent from the quenched KPZ (1/3) and EW (1/4) equations, and that is the meaning of our terminology "anomaly".

VI. SPECTRUM OF GROWTH RATES

Finally, in this section, we discuss the effect of the growth-rate spectrum on our model study.

The configurational property $\mathcal{A}$ is well defined even if we consider a continuous spectrum of growth rates α. Also, if we take the weighted Kuramoto oscillator's order parameters, $\mathcal{R}$, is still defined in the 1D case. The presence or absence of bimodality depends heavily on the proportion of the population corresponding to the growth rates (because even if the income levels of the distributions are significantly different, the height difference may be so large that the smaller distribution appears to be absorbed by the larger distribution). Hence, the second era of bimodality in global income inequality can be seen as a result of an appropriate distribution of income levels and population shares between developed and developing countries. However, the important thing is the regional segregation of income level is consistent with our result even if we take continuous spectrum of growth rates such as uniform distribution $\alpha_n \in [\alpha-\Delta\alpha, \alpha+\Delta\alpha]$ or normal distribution $\alpha_n \sim \mathcal{N}(\mu, \sigma^2)$ and it matches the concept of *location-based inequality* [8].

Since normally distributed growth rates in the 1D ring case has been tested in our previous study [9], it was intentionally omitted. Instead, we focus on clarifying the analytics of the binary case in detail. The analysis of the aggregate growth rate, Gini index, and correlations between growth rate and log-income through the configurational property $\mathcal{A}$ under a normally distributed growth rate, see the Supplemental Material (SM) of our previous paper [9].

[1] M. Medo, Breakdown of the mean-field approximation in a wealth distribution model, Journal of Statistical Mechanics: Theory and Experiment **2009**, P02014 (2009).

[2] T. Ma, J. G. Holden, and R. Serota, Distribution of wealth in a network model of the economy, Physica A: Statistical Mechanics and its Applications **392**, 2434 (2013).

[3] J.-P. Bouchaud, The self-organized criticality paradigm in economics & finance, arXiv preprint arXiv:2407.10284 (2024).

[4] K. I. Park, M. Park, *et al.*, *Fundamentals of probability and stochastic processes with applications to communications* (Springer, 2018).

[5] J.-P. Bouchaud and M. Mézard, Wealth condensation in a simple model of economy, Physica A: Statistical Mechanics and its Applications **282**, 536 (2000).

[6] W. Souma, Y. Fujiwara, and H. Aoyama, Small-world effects in wealth distribution, arXiv preprint cond-mat/0108482 (2001).

[7] B. Milanovic, The three eras of global inequality, 1820–2020 with the focus on the past thirty years, World Development **177**, 106516 (2024).

[8] B. Milanovic, Global inequality of opportunity: How much of our income is determined by where we live?, Review of Economics and Statistics **97**, 452 (2015).

[9] J. Hur, M. Ha, and H. Jeong, Interplay of network structure and talent configuration on wealth dynamics, Physical Review E **110**, 024312 (2024).