



# Power Laws in Stochastic Processes for Social Phenomena: An Introductory Review

Kumamoto, Shin-Ichiro  
Kamihigashi, Takashi

---

(Citation)

Frontiers in Physics, 6:20-20

(Issue Date)

2018-03-15

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

© 2018 Kumamoto and Kamihigashi.

This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and...

(URL)

<https://hdl.handle.net/20.500.14094/90004813>





# Power Laws in Stochastic Processes for Social Phenomena: An Introductory Review

Shin-Ichiro Kumamoto\* and Takashi Kamihigashi

Research Institute for Economics and Business Administration, Kobe University, Kobe, Japan

Many phenomena with power laws have been observed in various fields of the natural and social sciences, and these power laws are often interpreted as the macro behaviors of systems that consist of micro units. In this paper, we review some basic mathematical mechanisms that are known to generate power laws. In particular, we focus on stochastic processes including the Yule process and the Simon process as well as some recent models. The main purpose of this paper is to explain the mathematical details of their mechanisms in a self-contained manner.

**Keywords:** power law, Zipf's law, Pareto's law, preferential attachment, geometric brownian motion

## OPEN ACCESS

### Edited by:

Isamu Okada,  
Sōka University, Japan

### Reviewed by:

Francisco Wellington Lima,  
Federal University of Piauí, Brazil  
Renaud Lambiotte,  
University of Oxford, United Kingdom

### \*Correspondence:

Shin-Ichiro Kumamoto  
kumamoto@rieb.kobe-u.ac.jp

### Specialty section:

This article was submitted to  
Interdisciplinary Physics,  
a section of the journal  
Frontiers in Physics

**Received:** 13 October 2017

**Accepted:** 14 February 2018

**Published:** 15 March 2018

### Citation:

Kumamoto S-I and Kamihigashi T  
(2018) Power Laws in Stochastic  
Processes for Social Phenomena: An  
Introductory Review.  
Front. Phys. 6:20.  
doi: 10.3389/fphy.2018.00020

## 1. INTRODUCTION

Many phenomena with power laws have been observed in various fields of the natural and social sciences: physics, biology, earth planetary science, computer science, economics, and so on. These power laws can be interpreted as the macro behaviors of the systems that consist of micro units (i.e., agents, individuals, particles, and so on). In other words, the ensemble of dynamics of these micro units is observed as the behavior of the whole system such as a power law<sup>1</sup>. To obtain a deep understanding of the phenomenon for the system, we must first observe the behavior on the macro side, then assume the stochastic dynamics on the micro side, and finally reveal the theoretical method connecting both sides. Thus, the mechanisms generating power laws have been studied as the second and final steps in the study of power laws.

Next, we mathematically define the power law. When the probability density function  $p(x)$  for a continuous random variable<sup>2</sup>  $\hat{X}$  is given by

$$p(x) = Cx^{-\alpha} \quad (x \geq x_{\min}), \quad (1)$$

we say that  $\hat{X}$  satisfies the power law. The exponent  $\alpha$  is called the exponent of power law,  $C$  is the normalization constant, and  $x_{\min}$  is the minimum value that  $x$  satisfies the power law. The power law is the only function satisfying the scale-free property [1]

$$p(bx) = f(b)p(x) \quad \text{for any } b. \quad (2)$$

Then we define the cumulative distribution function  $P_{>}(x)$  as

<sup>1</sup>For the example of a city, the micro dynamics correspond to immigration, emigration, births, and deaths, and the macro behavior is the distribution of the population.

<sup>2</sup>The hat of  $\hat{O}$  means that  $\hat{O}$  is a random variable.

$$P_{>}(x) := P\{\hat{X} \geq x\} = \int_x^{\infty} p(x)dx. \quad (3)$$

When the probability density function satisfies the power law  $p(x) = Cx^{-\alpha}$ ,

$$P_{>}(x) \propto x^{-\alpha+1}. \quad (4)$$

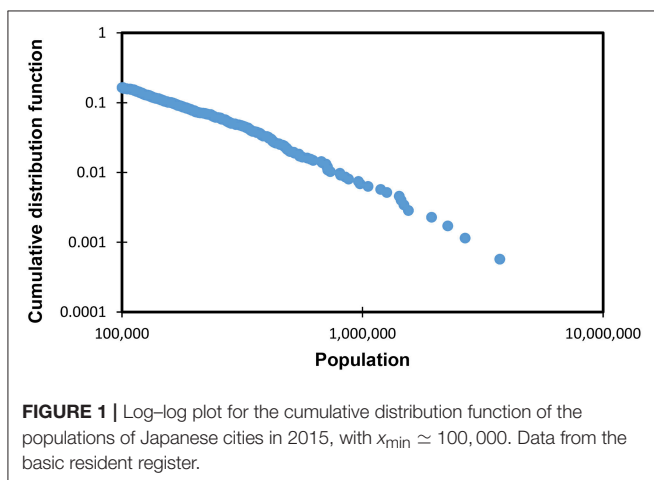
The behavior of the cumulative distribution function with the power law is a straight line in a log-log plot for  $x \geq x_{\min}$  (Figure 1).

Next, we list some examples of power laws in various phenomena.

- (a) Populations of cities [2].
- (b) Frequency of use of words [3, 4].
- (c) Number of papers published by scientists [5].
- (d) Number of citations received by papers [6].
- (e) Number of species in biological genera [7, 8].
- (f) Number of links on the World Wide Web [9].
- (g) Individual wealth and income [10].
- (h) Sizes of firms (the number of employees, assets, or market capitalization) [11–17].
- (i) Sizes of earthquakes [18].
- (j) Sizes of forest fires [19].

Furthermore, we partly list the generating mechanisms that are important for applications, and the phenomena to which they are applied in the above list, such as “mechanism  $\Rightarrow$  phenomena.”

- Growth and preferential attachment:
  - Yule process [20]  $\Rightarrow$  (e);
  - Simon process [21]  $\Rightarrow$  (a), (b), (c), (e), and (g);
  - Barabási–Albert (BA) model [22]  $\Rightarrow$  (d) and (f).
- Stochastic models based on Geometric Brownian motion (GBM):
  - GBM with a reflecting wall [23]  $\Rightarrow$  (a), (g), and (h);
  - GBM with reset events [24, 25]  $\Rightarrow$  (g).
  - Kesten process [26]  $\Rightarrow$  (g).



- Generalized Lotka–Volterra (GLV) model [27–29]  $\Rightarrow$  (g);
- Bouchaud–Mézard (BM) model [30]  $\Rightarrow$  (g).

- Combination of exponentials (change of variable) [31]  $\Rightarrow$  (b).
- Self-organized criticality [32]  $\Rightarrow$  (i) and (j).
- Highly optimized tolerance [33, 34]  $\Rightarrow$  (j).

Though there are many other generating mechanisms besides them<sup>3</sup>, the mechanisms of the above list are particularly well known and widely applied to phenomena in various fields.

In this paper, we focus on the generating mechanisms with the stochastic processes in the above list<sup>4</sup>: the growth and preferential attachment and the stochastic models based on the GBM, which, in particular, are widely applied in social science. In addition, we explain about the combination of exponentials that is related to the mechanism of the Yule process. We mainly give full details of the mathematical formalisms for these mechanisms in self-contained manner, because understanding them is important for researchers in any field to create new models generating power laws in empirical data. The necessary mathematical supplements to understand these mechanisms are given in the Appendix at the end of this paper.

## 2. GROWTH AND PREFERENTIAL ATTACHMENT

As the name suggests, this mechanism consists of the two characteristics: growth and preferential attachment. In the example of a city, the meanings of growth and preferential attachment are as follows.

- Growth: The number of cities increases.
- Preferential attachment: The more populated cities become, the higher the probability that the population will increase. Namely, it is “the rich get richer” process<sup>5</sup>.

In this section, we deal with the Yule process, the Simon process, and the BA model, which all have these two characteristics. The Yule process generates the power law about the number of species within genera in biology. The Simon process generates the power laws about the frequency of use of a word in a text, the population of cities, and so on (see the list in the Introduction for details). The BA model generates the power law about the number of edges incident to nodes in the network. We now explain in detail how these three mechanisms mathematically generate the power laws.

<sup>3</sup>Readers interested in more phenomena with power laws and their generating mechanisms should refer to the reviews and textbooks by Mitzenmacher [35], Newman [1], Sornette [36], Hayashi et al. [37], Farmer and Geanakoplos [38], Gabaix [39, 40], Simkin and Roychowdhury [41], Pinto et al. [42], Piantadosi [43], Machado et al. [44], and Slanina [45].

<sup>4</sup>Though the multiplicative process [46] is also the stochastic process, it is not explained in this paper because the multiplicative process is interpreted as the discrete-time version of the GBM of the continuous-time stochastic process. Namely, the multiplicative process is essentially equivalent to the GBM (see Appendix A.4).

<sup>5</sup>Preferential attachment is also called the Matthew effect [47] or the cumulative advantage [48].

## 2.1. Yule Process

The Yule process [20] was invented to model stochastic population growth with the preferential attachment process for the model of speciation in biology. In this process, new species and genera are born by biological mutations that are interpreted as the branchings from the lines of existing species in the evolutionary tree (Figures 2, 3).

These branchings occur as Poisson processes and add lines of new genera or species to the evolutionary tree. The Yule process mathematically corresponds to the stochastic process that the numbers of genera and species increase independently by following the linear birth processes (see Appendix A.3)<sup>6</sup>. In other words, we consider the evolutionary tree of species (Figure 2) and that of genera (Figure 3) separately.

In short, the Yule process is the combination of the stochastic processes for the numbers of species and genera (Figure 4) [41, 49, 50].

- The number of species within a genus increases as the linear birth process with the Poisson rate  $\lambda_s n_s$ , where  $\lambda_s$  is a positive constant and  $n_s$  is the number of species within the genus at that time.
- The number of genera within a family increases as the linear birth process with the Poisson rate  $\lambda_g n_g$ , where  $\lambda_g$  is a positive

constant and  $n_g$  is the number of genera within the family at that time.

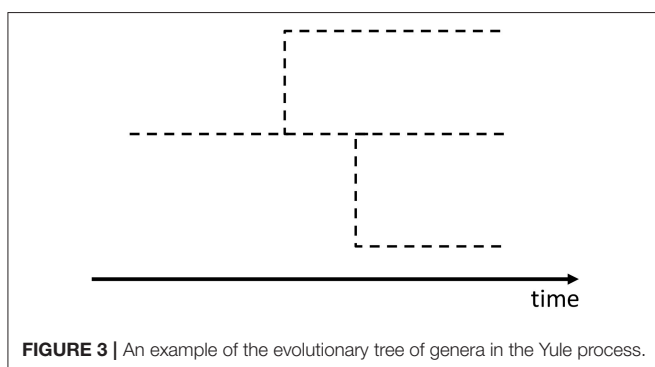
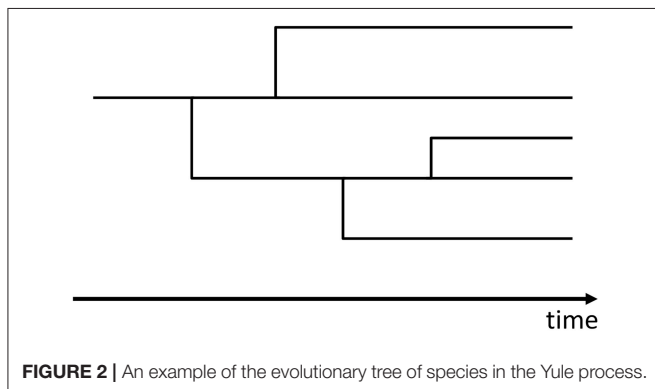
To obtain the probability distribution of the number of species within genera at a large time<sup>7</sup>, we need the conditional probability distribution of the number of species included in the genus whose age (i.e., the time intervals elapsed since the birth) is  $t$ . Let us use  $r_s(n, t)$  to denote its conditional probability distribution, where  $n \in \mathbb{N}$  is the number of species and  $t \in \mathbb{R}$  is the age of the genus.

First,  $r_s(1, t)$  is equivalent to the probability that no new species is born in  $(a, a + t]$  after the genus is born<sup>8</sup> at an arbitrary time  $a$ . Accordingly, we obtain  $r_s(1, t)$  from (A.2) as

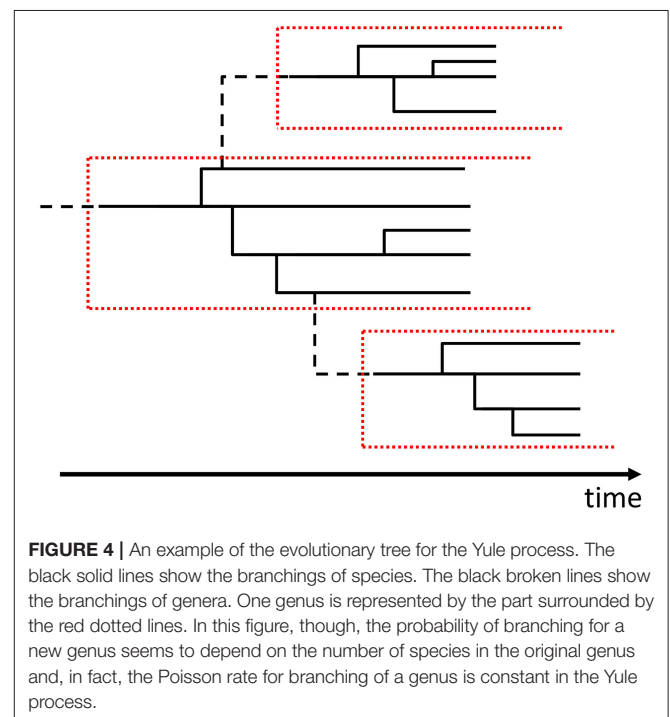
$$r_s(1, t) = P\{\hat{N}_s(a + t) - \hat{N}_s(a) = 0; \text{rate } \lambda_s\} = e^{-\lambda_s t}, \quad (5)$$

where  $\hat{N}_s(t)$  is the number of species born in  $(0, t]$  by the Poisson process with the Poisson rate  $\lambda_s$ .

Second, we calculate  $r_s(2, t)$ . It is equivalent to the probability that one new species is born in  $(a, a + t]$  after the genus is born at an arbitrary time  $a$ . Then we assume that one new species is born in the infinitesimal time interval  $[a + \tau_1, a + \tau_1 + d\tau_1]$ . From (A.2) and (A.3), we obtain the probabilities for one



<sup>6</sup>The characteristic of growth is the increase in the number of genera. The characteristic of preferential attachment is that the more species within a genus, the more new species are born.



<sup>7</sup>We consider the probability distribution only at a large time for the stationary state.

<sup>8</sup>This new genus is equivalent to the first species born in its own genus. Therefore, the new genus is counted as one for the number of species.

birth or no birth in each of the three divided time intervals: where the beta function  $B(a, b)$  is defined as

$$\begin{cases} P\{\text{no birth in } (a, a + \tau_1); \text{ rate } \lambda_s\} = e^{-\lambda_s \tau_1}, \\ P\{\text{one birth in } [a + \tau_1, a + \tau_1 + d\tau_1]; \text{ rate } \lambda_s\} = P\{\hat{N}_s(a + \tau_1 + d\tau_1) - \hat{N}_s(a + \tau_1) = 1; \text{ rate } \lambda_s\} \\ \quad = e^{-\lambda_s d\tau_1} \lambda_s d\tau_1 \simeq \lambda_s d\tau_1, \\ P\{\text{no birth in } [a + \tau_1 + d\tau_1, a + t]; \text{ rate } 2\lambda_s\} = e^{-2\lambda_s(a+t-\tau_1-d\tau_1)} \simeq e^{-2\lambda_s(a+t-\tau_1)}. \end{cases} \quad (6)$$

Integrating the product of these probabilities with respect to  $\tau_1$ , we obtain

$$r_s(2, t) = \int_0^t e^{-\lambda_s \tau_1} \lambda_s e^{-2\lambda_s(t-\tau_1)} d\tau_1 = e^{-\lambda_s t} (1 - e^{-\lambda_s t}). \quad (7)$$

Similarly,  $r_s(3, t)$  is given by

$$\begin{aligned} r_s(3, t) &= \int_0^t e^{-\lambda_s \tau_1} \lambda_s d\tau_1 \int_{\tau_1}^t e^{-2\lambda_s(\tau_2-\tau_1)} (2\lambda_s) e^{-3\lambda_s(t-\tau_2)} d\tau_2 \\ &= e^{-\lambda_s t} (1 - e^{-\lambda_s t})^2. \end{aligned} \quad (8)$$

Finally, repeating the same procedure, we obtain  $r_s(n, t)$ , that is, the conditional probability distribution of the number of species included in the genus at the age of  $t$ :

$$\begin{aligned} r_s(n, t) &= e^{-n\lambda_s t} \prod_{k=1}^{n-1} \left[ \int_{\tau_{k-1}}^t e^{\lambda_s \tau_k} k \lambda_s d\tau_k \right] \quad (\tau_0 := 0) \\ &= e^{-n\lambda_s t} (n-1)! \prod_{k=1}^{n-1} \left[ \int_{x_{k-1}}^{x_k} dx_k \right] \quad (x_k := e^{\lambda_s \tau_k}, x_0 := 1) \\ &= e^{-\lambda_s t} (1 - e^{-\lambda_s t})^{n-1}. \end{aligned} \quad (9)$$

Next, let  $\ell_g(t)$  be the probability distribution function for the age of genera at a large time in the linear birth process. It is given by (A.15) as

$$\ell_g(t) = \lambda_g e^{-\lambda_g t}. \quad (10)$$

Consequently, the probability density of the number of species within genera at a large time, denoted by  $q(n)$ , is given by integrating the product of the conditional probability distribution of the number of species within genera and the probability density function for the age of genera at a large time:

$$\begin{aligned} q(n) &= \int_0^\infty r_s(n, t) \ell_g(t) dt = \int_0^\infty e^{-\lambda_s t} (1 - e^{-\lambda_s t})^{n-1} \lambda_g e^{-\lambda_g t} dt \\ &= \frac{\lambda_g}{\lambda_s} \int_0^1 x^{\frac{\lambda_g}{\lambda_s}} (1-x)^{n-1} dx \quad (x := e^{-\lambda_s t}) \\ &= : \frac{\lambda_g}{\lambda_s} B\left(\frac{\lambda_g}{\lambda_s} + 1, n\right), \end{aligned} \quad (11)$$

$$\begin{aligned} B(a, b) &:= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ \left( \Gamma(a) &:= \int_0^\infty t^{a-1} e^{-t} dt \right). \end{aligned} \quad (12)$$

When  $b$  takes a large value, the beta function is approximately

$$B(a, b) \propto b^{-a} \quad (b \gg 1). \quad (13)$$

Therefore, for a large number of species, the probability distribution of the number of species within genera at a large time satisfies the power law as

$$q(n) \propto n^{-\left(\frac{\lambda_g}{\lambda_s} + 1\right)} \quad (n \gg 1), \quad (14)$$

where the exponent of power law is  $\frac{\lambda_g}{\lambda_s} + 1$ .

## 2.2. Simon Process

The Simon process [21] is interpreted as a discrete-time stochastic process for the growth in the numbers of urns and balls contained in those urns: an urn and the number of balls in the urn correspond to a word and the number of times that the word is used. In this stochastic process, a certain number of balls are newly added and stochastically distributed to the existing urns containing some balls at each time step. After that, one urn containing a certain number of balls (it need not be the same as the number of balls added above) is also added newly. Repeating this procedure, the number of balls and urns grows stochastically.

We calculate the stationary probability distribution of balls contained in urns at a large time.

First, we define all quantities for the Simon process by using the following notation:

- $t$  ( $= 0, 1, 2, \dots$ ), discrete time;
- $k_0$ , number of balls contained in each urn in the initial state (before balls are added);
- $m$ , number of balls added at each time step;
- $B(t)$  ( $= B(0) + t(m + k_0)$ ), total number of balls before balls are distributed at  $t$ ;
- $U(t)$  ( $= U(0) + t$ ), number of urns before balls are distributed at  $t$ ;
- group- $(k)$ , group of all the urns containing  $k$  balls;
- $\hat{f}(k, t)$ , number of urns belonging to the group- $(k)$  before balls are distributed at  $t$ .

Next, we provide the detailed procedure with the stochastic rule as follows (Figure 5).

1. There are  $U(0)$  urns containing  $k_0$  balls at the initial time<sup>9</sup>  $t = 0$ .
2. The  $m$  balls are newly added at each time step<sup>10</sup>.
3. Each of the  $m$  balls is distributed once for each group- $(k)$  with the probability  $\frac{k\hat{f}(k, t)}{B(t)}$ <sup>11</sup>.
4. Then the balls distributed to the group- $(k)$  are further distributed to the urns within the group with arbitrary probabilities<sup>12</sup> with the assumption that each urn can only get up to one ball at each time step<sup>13</sup>.
5. At the end of each time step, one urn containing  $k_0$  balls is added.
6. We repeat steps 2–5.

Then we can obtain the expectation values of  $E[\hat{f}(k, t + 1)]$  ( $k \geq k_0$ ) from the above stochastic rule as

$$\begin{cases} E[\hat{f}(k, t + 1)] = \hat{f}(k, t) - \frac{mk\hat{f}(k, t)}{B(t)} + \frac{m(k-1)\hat{f}(k-1, t)}{B(t)} \\ \quad (k > k_0), \\ E[\hat{f}(k_0, t + 1)] = \hat{f}(k_0, t) - \frac{mk_0\hat{f}(k_0, t)}{B(t)} + 1. \end{cases} \quad (15)$$

At a large time  $t$ , we can make an approximation  $\hat{f}(k, t) \simeq E[\hat{f}(k, t)]$  for  $k \geq k_0$  and obtain

$$\begin{cases} E[\hat{f}(k, t + 1)] \simeq E[\hat{f}(k, t)] - \frac{mkE[\hat{f}(k, t)]}{B(t)} + \frac{m(k-1)E[\hat{f}(k-1, t)]}{B(t)} \\ \quad (k > k_0) \\ E[\hat{f}(k_0, t + 1)] \simeq E[\hat{f}(k_0, t)] - \frac{mk_0E[\hat{f}(k_0, t)]}{B(t)} + 1. \end{cases} \quad (16)$$

The probability distribution of the number of balls in urns, denoted by  $p(k, t)$ , can be represented by  $E[\hat{f}(k, t + 1)]$ :

$$p(k, t) = \frac{E[\hat{f}(k, t)]}{U(t)}. \quad (17)$$

<sup>9</sup>Since we finally take the limit  $t \rightarrow \infty$ , the initial state does not actually affect the stationary state. However, to make it easier to imagine the procedure, we set the initial state in this manner.

<sup>10</sup>This shows the characteristic of growth.

<sup>11</sup>This shows the characteristic of preferential attachment.

<sup>12</sup>When distributing balls in the group- $(k)$ , we do not set the probability that each urn in the group gets one ball. To obtain a master equation later, we only have to know the number of the balls distributed to the group- $(k)$  under the condition that each urn can only get one ball at most. Namely, setting those probabilities is equivalent to imposing too strong a condition to obtain the master equation.

<sup>13</sup>Though one urn can get two or more balls, this possibility is small enough in the limit of large time. This is because the number of urns is large enough in a large time so that this possibility is ignored. Similarly, though more balls can be distributed than the number of urns in a group- $(k)$ , this possibility is also small enough in the limit of large time.

Consequently, the master equation for  $p(k, t)$  is given by

$$\begin{cases} U(t+1)p(k, t+1) = U(t)p(k, t) - \frac{mkU(t)}{B(t)}p(k, t) + \frac{m(k-1)U(t)}{B(t)}p(k-1, t) \quad (k > k_0), \\ U(t+1)p(k_0, t+1) = U(t)p(k_0, t) - \frac{mkU(t)}{B(t)}p(k_0, t) + 1. \end{cases} \quad (18)$$

We are interested in only the stationary distribution function  $p(k)$  that is defined as  $p(k, t)$  in the limit of large time:

$$p(k) := \lim_{t \rightarrow \infty} p(k, t). \quad (19)$$

Then, considering

$$\lim_{t \rightarrow \infty} \frac{U(t)}{B(t)} = \frac{1}{m + k_0} \quad (20)$$

and taking the limit  $t \rightarrow \infty$  for Equation (18), we obtain

$$\begin{cases} p(k) = \frac{k-1}{k+1+\frac{k_0}{m}}p(k-1) \quad (k > k_0), \\ p(k_0) = \frac{m+k_0}{k_0(m+1)+m}. \end{cases} \quad (21)$$

We can solve these equations recursively:

$$\begin{aligned} p(k) &= \frac{(k-1)(k-2) \cdots k_0}{\left(k+1+\frac{k_0}{m}\right)\left(k+\frac{k_0}{m}\right) \cdots \left(k_0+2+\frac{k_0}{m}\right)} p(k_0) \\ &= \frac{(k-1)(k-2) \cdots k_0}{(k-1+\alpha)(k-2+\alpha) \cdots (k_0+\alpha)} p(k_0) \quad \left(\alpha := 2 + \frac{k_0}{m}\right) \\ &= \frac{\Gamma(k)\Gamma(k_0+\alpha)}{\Gamma(k_0)\Gamma(k+\alpha)} p(k_0) \\ &= \frac{B(k, \alpha)}{B(k_0, \alpha)} p(k_0). \end{aligned} \quad (22)$$

For the large  $k$ , the stationary probability distribution of the number of balls in urns satisfies the power law as

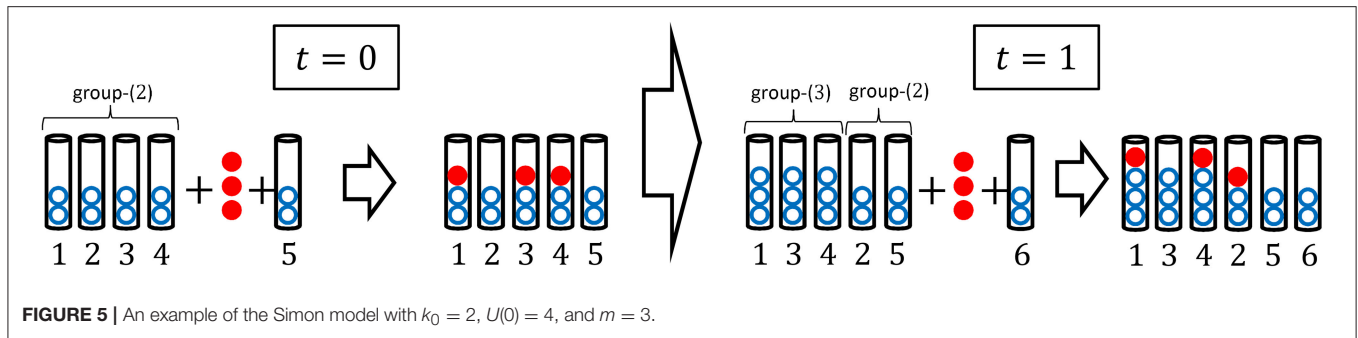
$$p(k) \propto k^{-\left(\frac{k_0}{m}+2\right)} \quad (k \gg 1), \quad (23)$$

where the exponent of power law is  $\frac{k_0}{m} + 2$ .

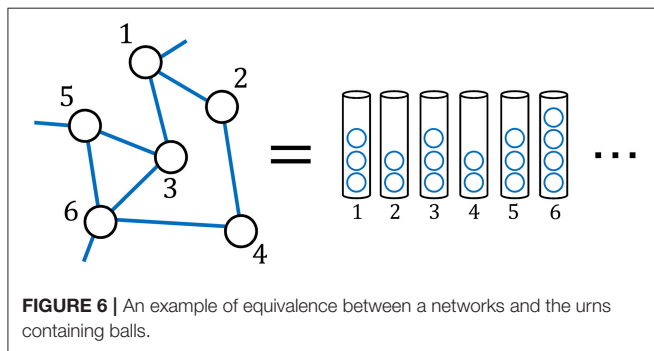
### 2.3. Barabási–Albert Model

The BA model [22] is one of the scale-free network models for the growth in the number of nodes and edges. Mathematically, the BA model can be interpreted as a special case of the Simon model. In particular, the nodes and edges in the BA model correspond to the urns and balls in the Simon model, respectively (**Figure 6**). In this model, one node with a certain number of edges are





**FIGURE 5** | An example of the Simon model with  $k_0 = 2$ ,  $U(0) = 4$ , and  $m = 3$ .



**FIGURE 6** | An example of equivalence between a networks and the urns containing balls.

newly added at each time step. Then following a stochastic rule, the edges of new node are connected to the existing nodes. Repeating this procedure, the number of nodes and edges grows stochastically.

We calculate the stationary probability distribution of edges connecting to nodes at a large time. First, we define all quantities for the BA model by using the following notation:

- $t (= 0, 1, 2, \dots)$ , discrete time;
- $k_0$ , number of edges that the additional new node has;
- $B(t) (= B(0) + 2tk_0)$ , total number of degrees in the network before the new node is added at  $t$ ;
- $U(t) (= U(0) + t)$ , number of nodes in the network before the new node is added at  $t$ ;
- $\hat{f}(k, t)$ , number of nodes with the degree  $k$  before the new node is added at  $t$ ;
- $\hat{k}_i(t)$ , number of the edges of node- $i$  (where  $i$  is the label of the node) before the new node is added at  $t$ .

Next, we give the detailed procedure with the stochastic rule for the BA model as follows (Figure 7).

1. At the initial time  $t = 0$ , there is an arbitrary connected network with  $U(0)$  nodes that are all connected to nodes other than themselves<sup>14</sup>.
2. One new node with  $k_0$  edges is added<sup>15</sup>.

<sup>14</sup>Since we finally take the limit  $t \rightarrow \infty$  as in the Simon model, the initial state does not actually affect the stationary state.

<sup>15</sup>This shows the characteristic of growth.

3. The  $k_0$  edges of the new node are connected to the existing nodes following the stochastic rule<sup>16,17</sup>: the probability that one edge is connected to the existing node- $i$  is  $\frac{\hat{k}_i(t)}{B(t)}$  under the assumption that each node can only connect to one node at each time step<sup>18</sup>.
4. We repeat steps 2 and 3.

Consequently, we obtain the same master equation for the probability distribution of edges as (18) with  $m = k_0$ :

$$\begin{cases} U(t+1)p(k, t+1) = U(t)p(k, t) - \frac{k_0 k U(t)}{B(t)} p(k, t) + \frac{k_0(k-1)U(t)}{B(t)} p(k-1, t) & (k > k_0), \\ U(t+1)p(k_0, t+1) = U(t)p(k_0, t) - \frac{k_0 k U(t)}{B(t)} p(k_0, t) + 1. \end{cases} \quad (24)$$

We can solve this master equation and obtain the stationary distribution function  $p(k) := \lim_{t \rightarrow \infty} p(k, t)$  for the large  $k$  from Equations (21–23):

$$p(k) \propto k^{-\alpha} \quad \left( \alpha := 2 + \frac{k_0}{k_0} = 3 \right), \quad (25)$$

where the exponent of power law is 3.

### 3. STOCHASTIC MODELS BASED ON GEOMETRIC BROWNIAN MOTION

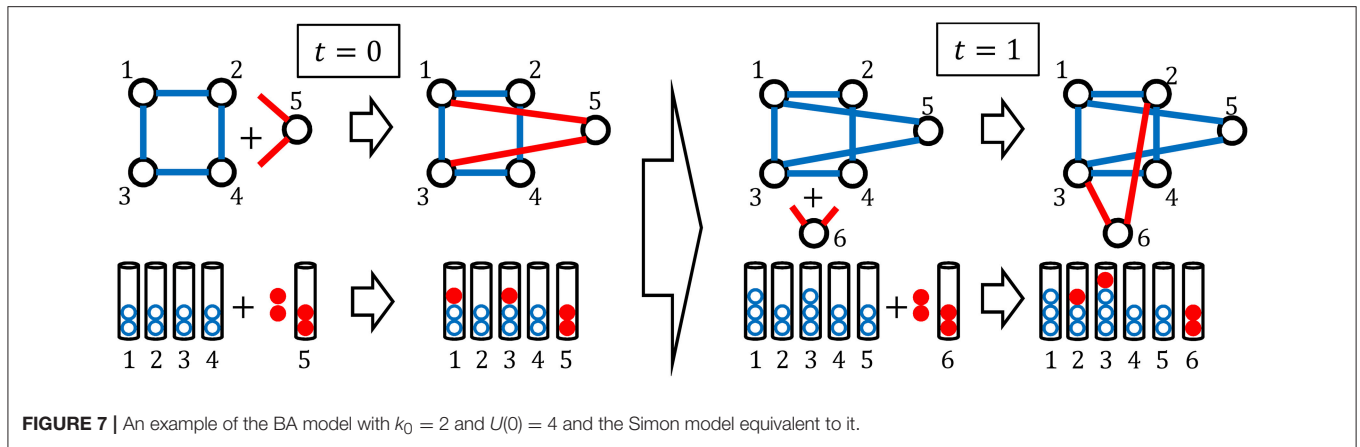
In this section we look at five stochastic processes, generating power laws, which can be represented by the stochastic differential equations (SDEs). They all are mathematically based on the GBM and accompanied by a constraint (i.e., additional condition) or additional terms to the SDE. The constraints correspond to a reflecting wall<sup>19</sup> as a boundary condition [23],

<sup>16</sup>This shows the characteristic of preferential attachment.

<sup>17</sup>This setting of probability is equivalent to the balls distributed to the group- $(k)$  being further distributed to the urns within the group with equal probabilities in step 4 in the Simon model. That is, the stochastic rule for the BA model is stronger than that of the Simon model as a condition.

<sup>18</sup>Though one node can actually connect two or more nodes, this possibility is small enough in the limit of large time. This is because the number of nodes is large enough in a large time so that this possibility is ignored.

<sup>19</sup>The reflecting wall means that there is the minimum value for a random variable (e.g., population of a city).



**FIGURE 7** | An example of the BA model with  $k_0 = 2$  and  $U(0) = 4$  and the Simon model equivalent to it.

and reset events (i.e., birth and death process<sup>20</sup>) [25]. The stochastic processes with additional terms to the SDE of GBM are the Kesten process, the GLV model, and the BM model. Though the effect of additional term to the GMB in the Kesten process is similar to a reflecting wall, those of the GLV model and BM model correspond to the interactions between particles, agents, or individuals. We mainly explain the mathematical formalisms and properties of these qualitatively different stochastic processes.

### 3.1. Geometric Brownian Motion

The GBM, on which many models for power laws are based, is one of the most important stochastic processes. It is mathematically defined by the SDE

$$d\hat{X}(t) = \mu \hat{X}(t)dt + \sigma \hat{X}(t)d\hat{B}(t), \quad (26)$$

where  $\hat{B}(t)$  is a standard Brownian motion,  $\mu$  is the drift, and  $\sigma$  is the volatility.

The SDE (Equation 26) gives us the partial differential equation (PDE), that is, the Fokker–Planck equation (FPE) [51]:

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x} \{ \mu x p(x, t) \} + \frac{\partial^2}{\partial x^2} \left\{ \frac{\sigma^2 x^2}{2} p(x, t) \right\}, \quad (27)$$

where  $p(x, t)$  is the probability density function. The solution of Equation (27) with the initial distribution  $p(x, 0) = \delta(x - x_0)$  is

$$p(x, t) = \frac{1}{x\sqrt{2\pi\sigma^2 t}} \exp \left[ -\frac{\left\{ \log x - \log x_0 - \left( \mu - \frac{\sigma^2}{2} \right) t \right\}^2}{2\sigma^2 t} \right], \quad (28)$$

where  $x_0$  is the initial position of the particle. This solution is the log-normal distribution where the expectation value and variance are

$$E[\hat{x}] = x_0 e^{\mu t}, \quad \text{Var}[\hat{x}] = x_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1). \quad (29)$$

In the limit  $t \rightarrow \infty$ , the log-normal distribution never converges to the stationary solution. To obtain it, therefore, we need to

<sup>20</sup>The birth and death process means that a new unit (e.g., city or firm) can be born at a rate and die at the same rate.

impose some additional conditions on the SDE (Equation 26) or modify the SDE itself. We introduce the conditions and modifications in the following sections.

### 3.2. GBM With a Reflecting Wall

We consider the GBM with the reflecting wall (see Appendix A.5 for details). The stationary solution  $p(x)$  for the FPE (Equation 27) is defined by

$$\frac{\partial p(x)}{\partial t} = 0, \quad (30)$$

which is equivalent to the second-order ordinary differential equation (ODE):

$$0 = -\frac{d}{dx} \{ \mu x p(x) \} + \frac{d^2}{dx^2} \left\{ \frac{\sigma^2 x^2}{2} p(x) \right\}. \quad (31)$$

As a result, we obtain the first-order ODE:

$$\mu x p(x) - \frac{d}{dx} \left\{ \frac{\sigma^2 x^2}{2} p(x) \right\} = D, \quad (32)$$

where  $D$  is an arbitrary constant. We take  $D = 0$  to obtain a normalizable power-law probability distribution. The solution of Equation (32) with  $D = 0$  is

$$p(x) = Cx^{-\alpha} \quad \left( C := p(x_0)x_0^\alpha, \quad \alpha := 2 - \frac{2\mu}{\sigma^2} \right), \quad (33)$$

where  $x_0$  is an arbitrary constant. For this stationary solution  $p(x)$  to exist, it must satisfy the normalization condition:

$$1 = \int_{x_{\min}}^{x_{\max}} p(x) dx. \quad (34)$$

We set the reflecting wall at  $x = x_{\min} (> 0)$  and take  $x_{\max} = \infty$ . The existence of the reflecting wall is mathematically equivalent to the conditions  $\hat{X}(t) > x_{\min}$  and  $p(x) = 0$  for  $x < x_{\min}$ . Then we assume  $\alpha > 1$ . The normalization condition

$$1 = \int_{x_{\min}}^{\infty} p(x) dx = \frac{C}{\alpha - 1} (x_{\min})^{-\alpha+1} \quad (35)$$



determines the constant  $C$  as

$$C = (\alpha - 1)(x_{\min})^{-\alpha+1}. \quad (36)$$

Thus, we have the normalized stationary solution

$$p(x) = (\alpha - 1)(x_{\min})^{-\alpha+1} x^{-\alpha} \quad \left( \alpha = 2 - \frac{2\mu}{\sigma^2} > 1 \right), \quad (37)$$

where the exponent of power law is  $2 - \frac{2\mu}{\sigma^2}$ .

Next, we generalize this formalism from the GBA to the Itô process which can have the stationary solution [52]:

$$d\hat{X}(t) = a(\hat{X}(t))dt + b(\hat{X}(t))d\hat{B}(t). \quad (38)$$

The stationary solution (see Appendix A.5 for details) is given by

$$p(x) = \frac{C}{b(x)^2} \exp \left[ \int_{x_0}^x \frac{2a(x')}{b(x')^2} dx' \right], \quad (39)$$

where  $C$  is the normalization constant. Following Yakovenko and Rosser [53] and Banerjee and Yakovenko [54], we take  $a(x)$  and  $b(x)$  as

$$a(x) = \mu x + \mu^*, \quad b(x) = \sigma \sqrt{2(x^2 + x^{*2})}, \quad (40)$$

which is interpreted as a kind of qualitative combination of the generalized Wiener process<sup>21</sup> and GBM. Consequently, we obtain the stationary solution

$$p(x) = C \left[ 1 + \left( \frac{x}{x^*} \right)^2 \right]^{\frac{\mu}{2\sigma^2}-1} \exp \left[ \frac{\mu^*}{\sigma^2 x^*} \arctan \left( \frac{x}{x^*} \right) \right]. \quad (41)$$

For  $x \ll x^*$ , the stationary solution becomes the exponential distribution while for the large  $x$ , it satisfies the power law as

$$p(x) \propto x^{-(2-\frac{\mu}{\sigma^2})} \quad (x \gg x^*), \quad (42)$$

where the exponent of power law is  $2 - \frac{\mu}{\sigma^2}$ .

### 3.3. GBM With Reset Events

We consider the particles that follow the GBM with the reset events, that is, the birth and death events<sup>22</sup>. For simplicity, we assume that particles can disappear with a certain probability by following a Poisson process and immediately appear at a point so that the number of particles is conserved. By these reset events, the FPE (Equation 27) is changed into

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} = & -\frac{\partial}{\partial x} \{ \mu x p(x, t) \} + \frac{\partial^2}{\partial x^2} \left\{ \frac{\sigma^2 x^2}{2} p(x, t) \right\} \\ & + \eta \delta(x - x^*) - \eta p(x, t), \end{aligned} \quad (43)$$

<sup>21</sup>The SDE of generalized Wiener process is represented by  $d\hat{X}(t) = a dt + b d\hat{B}(t)$ , where  $a$  and  $b$  are constants.

<sup>22</sup>Following Gabaix [39] and Toda [55], we derive the stationary probability density function.

where  $\eta$  is the probability for a particle in  $[x, x + dx)$  to disappear per the time interval  $dt$ , and the particle reappears immediately at  $x = x^* (> 0)$ . Accordingly, we obtain the second-order ODE for the stationary solution  $p(x)$ :

$$0 = -\frac{d}{dx} \{ \mu x p(x) \} + \frac{d^2}{dx^2} \left\{ \frac{(\sigma x)^2}{2} p(x) \right\} - \eta p(x), \quad (44)$$

which is held except for  $x = x^*$ . To solve this equation easily, we change the variable  $x$  into  $y := \log x$ . The new probability density function  $q(y)$  is determined by

$$q(y) = p(x) \left| \frac{dx}{dy} \right|. \quad (45)$$

Then we obtain the ODE for  $q(y)$ :

$$0 = -\left( \mu - \frac{\sigma^2}{2} \right) \frac{dq(y)}{dy} + \frac{\sigma^2}{2} \frac{d^2 q(y)}{dy^2} - \eta q(y), \quad (46)$$

except for  $y = y^*$  ( $y^* := \log x^*$ ). The general solution of this second-order ODE is

$$\begin{cases} q(y) = C_1 e^{\lambda_1 y} + C_2 e^{\lambda_2 y}, \\ \lambda_1 = \frac{1}{\sigma^2} \left( \mu - \frac{\sigma^2}{2} + \sqrt{\left( \mu - \frac{\sigma^2}{2} \right)^2 + 2\sigma^2 \eta} \right) > 0, \\ \lambda_2 = \frac{1}{\sigma^2} \left( \mu - \frac{\sigma^2}{2} - \sqrt{\left( \mu - \frac{\sigma^2}{2} \right)^2 + 2\sigma^2 \eta} \right) < 0, \end{cases} \quad (47)$$

where  $C_1$  and  $C_2$  are the arbitrary constants determined by the normalization condition:

$$1 = \int_0^\infty p(x) dx = \int_{-\infty}^\infty q(y) dy. \quad (48)$$

To normalize the solution (Equation 47), we impose the boundary conditions  $q(\infty) = q(-\infty) = 0$ , which result in  $C_1 = 0$  for  $y \geq y^*$  and  $C_2 = 0$  for  $y < y^*$ , that is,

$$q(y) = \begin{cases} C_1 e^{\lambda_1 y} & (y < y^*), \\ C_2 e^{\lambda_2 y} & (y \geq y^*). \end{cases} \quad (49)$$

Accordingly, the normalization condition

$$1 = \int_{-\infty}^{y^*} C_1 e^{\lambda_1 y} dy + \int_{y^*}^\infty C_2 e^{\lambda_2 y} dy \quad (50)$$

and the continuous condition at  $y = y^*$ , namely,  $C_1 e^{\lambda_1 y^*} = C_2 e^{\lambda_2 y^*}$  give us the normalized solution of Equation (46) as

$$q(y) = \begin{cases} \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{\lambda_1(y-y^*)} & (y < y^*), \\ \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{\lambda_2(y-y^*)} & (y \geq y^*). \end{cases} \quad (51)$$

Consequently, we obtain the solution of Equation (44):

$$p(x) = \frac{q(\log x)}{x} = \begin{cases} \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} (x^*)^{-\lambda_1} x^{\lambda_1 - 1} & (0 < x < x^*), \\ \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} (x^*)^{-\lambda_2} x^{\lambda_2 - 1} & (x^* \leq x), \end{cases} \quad (52)$$

which is called the double Pareto distribution [25]. The exponents of the power law are  $1 - \lambda_1$  and  $1 - \lambda_2$ .

Next, we derive the probability density function (Equation 52) by another method as follows [56]. The lifetimes of particles are independently distributed with the exponential distribution as  $\ell_{LT}(\tau) = \eta e^{-\eta\tau}$ , because the death events occur as a Poisson process, with rate  $\eta$ , which have the time-reversal symmetry property. Accordingly, the ages of particles (i.e., the time intervals elapsed since the birth of them) at a large time are also independently distributed with the exponential distribution:

$$\ell_A(t) = \eta e^{-\eta t}. \quad (53)$$

The probability density function of particles of age  $t$  as the conditional probability distribution is given by the log-normal distribution (Equation 28) with  $x_0 = x^*$ . Consequently, the probability density function of the coordinate of particle at a large time, denoted by  $p(x)$ , is given by integrating the product of Equations (53) and (28):

$$p(x) = \int_0^\infty \eta e^{-\eta t} \frac{1}{x\sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{\left\{\log x - \log x^* - \left(\mu - \frac{\sigma^2}{2}\right)t\right\}^2}{2\sigma^2 t}\right] dt. \quad (54)$$

We can calculate this with the change of variable  $u^2 := t$  and the identity [35]

$$\int_0^\infty \exp\left(-au^2 - \frac{b}{u^2}\right) du = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp(-2\sqrt{ab}). \quad (55)$$

Thus, we obtain the same result with Equation (52)<sup>23</sup> without solving the ODE (Equation 44).

### 3.4. Kesten Process

The Kesten process [26] is defined as a stochastic process whereby an additional term is added to the SDE of the GBM; namely, the SDE is represented by

$$d\hat{X}(t) = \mu\hat{X}(t)dt + \sigma\hat{X}(t)d\hat{B}(t) + \hat{c}dt, \quad (56)$$

where  $\hat{c}$ , in the additional term, is a random variable. We can expect that the additional term prevents  $\hat{X}(t)$  from decreasing toward  $-\infty$  in a similar way as the reflecting wall in section 3.2

<sup>23</sup>The two solutions in Equation (52) result from  $\sqrt{\left(\log \frac{x}{x^*}\right)^2} = -\log \frac{x}{x^*}$  for  $(0 < x < x^*)$ , and  $\sqrt{\left(\log \frac{x}{x^*}\right)^2} = \log \frac{x}{x^*}$  for  $(x^* \leq x)$ .

Here, we simply take  $\hat{c}$  as a positive constant:  $\hat{c} = c (> 0)$ . We then obtain the FPE for the probability density function:

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x} \{(\mu x + c)p(x, t)\} + \frac{\partial^2}{\partial x^2} \left\{ \frac{\sigma^2 x^2}{2} p(x, t) \right\}. \quad (57)$$

The ODE for the stationary solution  $p(x)$  is given by

$$0 = -\frac{d}{dx} \{(\mu x + c)p(x)\} + \frac{d^2}{dx^2} \left\{ \frac{(\sigma x)^2}{2} p(x) \right\}. \quad (58)$$

Consequently, we obtain the normalized stationary solution of Equation (58)<sup>24</sup>:

$$p(x) = \frac{1}{\Gamma(\alpha - 1)} \left( \frac{2c}{\sigma^2} \right)^{\alpha - 1} \exp\left[-\frac{2c}{\sigma^2 x}\right] x^{-\alpha} \left( \alpha := 2 - \frac{2\mu}{\sigma^2} \right), \quad (59)$$

where  $\Gamma(\alpha)$  is the gamma function defined in Equation (12). For the large  $x$ , the stationary solution satisfies the power law given as

$$p(x) \propto x^{-\beta} \quad (x \gg 1), \quad (60)$$

where the exponent of the power law is  $2 - \frac{2\mu}{\sigma^2}$ . Although  $c$ , in the additional term, achieves the stationary state, it is independent of the exponent. It is worth noting that the exponent of the power law is affected not by the constant  $c$  of the additional term, but by the drift  $\mu$  and volatility  $\sigma$  of the GBM. The additional term affects only the lower tail of the probability density function. Even for  $c$  as a random variable, these properties are invariant.

### 3.5. Generalized Lotka–Volterra Model

The GLV model was introduced for the analysis of individual income distribution. We consider the dynamical system composed of  $N$  agents (individuals) with incomes that grow by the GBM process and have the interactions for the redistribution of wealth [27–29]. The GLV model is represented by the system of SDEs called the GLV equations:

$$d\hat{X}_i(t) = \mu\hat{X}_i(t)dt + \sigma\hat{X}_i(t)d\hat{B}_i(t) + \xi\hat{U}(t)dt - \eta\hat{U}(t)\hat{X}_i(t)dt \\ \left( \hat{U}(t) := \frac{1}{N} \sum_{i=1}^N \hat{X}_i(t), \xi > 0 \right), \quad (61)$$

where  $\hat{X}_i(t)$  is the individual income of agent  $i$  ( $i = 1, 2, \dots, N$ ) at  $t$ , and  $\hat{U}(t)$  is the average income for the whole system. To keep  $\hat{X}_i(t) > 0$ , the third term in RHS of Equation (61) redistributes a fraction of the total income for the whole system. This term can be interpreted as the effect of a tax or social security policy. The fourth term controls the growth of whole system and can be interpreted as the effect of finiteness of resources, technological innovations, wars, natural disasters, and so on.

The GLV equations have no stationary solution, and the total income for the entire system is not constant with time. Here, we introduce the new random variable as the relative value:

$$\hat{Y}_i(t) := \frac{\hat{X}_i(t)}{\hat{U}(t)}. \quad (62)$$

<sup>24</sup>Following Slanina [45], we solve the ODE.

Then we obtain the SDEs for  $\hat{Y}_i(t)$  as

$$\begin{aligned} d\hat{Y}_i(t) &= \frac{d\hat{X}_i(t)}{\hat{U}(t)} - \frac{\hat{X}_i(t)d\hat{U}(t)}{\hat{U}(t)^2} \\ &= \xi(1 - \hat{Y}_i(t))dt + \sigma \hat{Y}_i(t)d\hat{B}_i(t) - \frac{\sigma \hat{Y}_i(t)}{N\hat{U}(t)} \sum_{i=1}^N \hat{X}_i(t)d\hat{B}_i(t), \end{aligned} \quad (63)$$

where the last term in the second row is of the order  $N^{-\frac{1}{2}}$ , because the standard deviation of  $\sum_{i=1}^N \hat{X}_i(t)d\hat{B}_i(t)$  is of the order  $\sqrt{N}$ .

We then take the large  $N$  limit as the mean field approximation and obtain the new system of SDEs:

$$d\hat{Y}_i(t) \simeq -\xi \hat{Y}_i(t)dt + \sigma \hat{Y}_i(t)d\hat{B}_i(t) + \xi dt, \quad (64)$$

which has the same form as the SDE of Equation (56) in the Kesten process. We can use the result of Equation (59) to obtain the normalized stationary probability density:

$$q(y) = \frac{1}{\Gamma(\alpha - 1)} \left( \frac{2\xi}{\sigma^2} \right)^{\alpha-1} \exp \left[ -\frac{2\xi}{\sigma^2 y} \right] y^{-\alpha} \left( \alpha := 2 + \frac{2\xi}{\sigma^2} \right). \quad (65)$$

For large  $y$ , the stationary solution satisfies the power law as follows:

$$q(y) \propto y^{-\alpha} \quad (y \gg 1), \quad (66)$$

where the exponent of the power law is  $2 + \frac{2\xi}{\sigma^2}$ . Consequently, by a change of variables, and the mean field approximation, the GLV model with interactions gives us the same result as that obtained by the Kesten process without interactions.

### 3.6. Bouchaud–Mézard Model

The BM model was introduced for the analysis of wealth distribution [30, 57–59]. We suppose there is an economic network composed of  $N$  agents (individuals) with wealth that grows by the GBM process and is redistributed by the exchanges between agents. The BM model is represented by the system of SDEs as follows:

$$d\hat{X}_i(t) = \mu \hat{X}_i(t)dt + \sigma \hat{X}_i(t)d\hat{B}_i(t) + \sum_{j(i \neq j)} a_{ij}(\hat{X}_j(t) - \hat{X}_i(t))dt \quad (67)$$

where  $\hat{X}_i(t)$  is the individual wealth of agent  $i$  at  $t$ , and  $a_{ij}$  is the positive exchange rate between agent  $i$  and  $j$ . The wealth is exchanged by the third term in RHS of Equation (67), which can be interpreted as a kind of trading in the economic network.

For simplicity, we take  $a_{ij}$  as the constant  $\frac{a}{N}$  ( $> 0$ ) in preparation for the mean field approximation. Here, we again introduce the new random variables as the average of wealth and the relative value:

$$\hat{Y}_i(t) := \frac{\hat{X}_i(t)}{\hat{U}(t)} \quad \left( \hat{U}(t) := \frac{1}{N} \sum_{i=1}^N \hat{X}_i(t) \right). \quad (68)$$

We then obtain the SDEs for  $\hat{Y}_i(t)$  in the mean field approximation:

$$\begin{aligned} d\hat{Y}_i(t) &= \frac{d\hat{X}_i(t)}{\hat{U}(t)} - \frac{\hat{X}_i(t)d\hat{U}(t)}{\hat{U}(t)^2} \\ &\simeq -a\hat{Y}_i(t)dt + \sigma \hat{Y}_i(t)d\hat{B}_i(t) + a dt, \end{aligned} \quad (69)$$

which has the same form as the SDE of Equation (64) in the LV model. Consequently, we obtain the normalized stationary solution:

$$q(y) = \frac{1}{\Gamma(\alpha - 1)} \left( \frac{2a}{\sigma^2} \right)^{\alpha-1} \exp \left[ -\frac{2a}{\sigma^2 y} \right] y^{-\alpha} \left( \alpha := 2 + \frac{2a}{\sigma^2} \right). \quad (70)$$

For large  $y$ , the stationary solution satisfies the following power law:

$$q(y) \propto y^{-\alpha} \quad (y \gg 1), \quad (71)$$

where the exponent of the power law is  $2 + \frac{2a}{\sigma^2}$ . It is worth noting that though the forms of the additional terms in the GLV model and BM model are quantitatively different from those of the Kesten process, the results are eventually the same in the mean field approximation.

## 4. COMBINATION OF EXPONENTIALS

When we have a probability density or distribution function for a random variable, we can obtain a new distribution by a change of variable. In particular, we can obtain a power law function from an exponential distribution by taking a new variable as the exponential function of the original variable. This mechanism was used to interpret the observed power law for the frequency of use of words with the assumption of random typings on a typewriter [31]. In this section, firstly we formalize this mechanism. Then we give the examples of applications to the Yule process and critical phenomena in physics.

### 4.1. General Formalism

Suppose the probability density function for a continuous random variable  $x$  is given by

$$p(x) = Ae^{ax} \quad (A > 0). \quad (72)$$

We change the variable  $x$  into the new variable  $y$  as

$$y = Be^{bx} \quad (B > 0). \quad (73)$$

Thus the new probability density function  $q(y)$  is obtained as

$$q(y) = p(x) \left| \frac{dx}{dy} \right| = \frac{A}{|b|B^{\frac{a}{b}}} y^{\frac{a}{b}-1} \propto y^{\frac{a}{b}-1}, \quad (74)$$

where the exponent of power law is  $\frac{a}{b} - 1$ .

Similarly, when the  $x$  is a discrete random variable, the new probability distribution function  $q(y)$  is obtained as

$$q(y) = p(x) = \frac{A}{B^{\frac{a}{b}}} y^{\frac{a}{b}} \propto y^{\frac{a}{b}}, \quad (75)$$

where the exponent of power law is  $\frac{a}{b}$ .

## 4.2. Application to Yule Process

The power law of the Yule process can be interpreted using a combination of exponentials with a rough approximation [41]. Firstly, by changing the Poisson rate  $\lambda_s$  into  $\lambda_g$  in (A.15), the probability density function of the age of genera at a large time is obtained as

$$p(t) = \lambda_g e^{-\lambda_g t}. \quad (76)$$

Then, from (A.12) with  $n_{s0} = 1$ , we approximately obtain the number of species within the genus of age  $t$  as

$$n(t) \simeq E[\hat{N}_s(t)] = e^{\lambda_s t}. \quad (77)$$

Finally, taking  $A = \lambda_g$ ,  $a = -\lambda_g$ ,  $B = 1$ , and  $b = \lambda_s$  in Equation (74), the probability density function of the number of species within genera is

$$q(n) = \frac{\lambda_g}{\lambda_s} n^{-\left(\frac{\lambda_g}{\lambda_s} + 1\right)}, \quad (78)$$

where the exponent of power law is  $\frac{\lambda_g}{\lambda_s} + 1$ . This exponent coincides with Equation (14). Thus the generating mechanism of power law in the Yule process can be roughly interpreted as a combination of exponentials as well.

## 4.3. Application to Critical Phenomena

It is well-known that in certain critical phenomena, some physical quantities (e.g., correlation length, susceptibility, and specific heat) take the form of power functions of the reduced temperature  $\frac{T-T_c}{T_c}$  near the critical temperature  $T_c$ . By the renormalization group analysis [60], this property can be interpreted as emerging from a combination of exponentials [41].

We consider two physical quantities  $x$  and  $y$  whose scaling dimensions are  $d_x$  and  $d_y$ , respectively. When we perform the scale transformation (i.e., renormalization group transformation) by the scaling factor  $b$  near the critical point, we suppose that  $x$  and  $y$  are multiplied by  $\lambda_x$  and  $\lambda_y$ , respectively. By the dimensional analysis, we obtain

$$\lambda_x = b^{d_x}, \quad \lambda_y = b^{d_y} \quad \left( \frac{\log \lambda_y}{\log \lambda_x} = \frac{d_y}{d_x} \right). \quad (79)$$

Then we obtain geometric progressions for the transformed  $x$  and  $y$ :

$$\begin{cases} x: x_0 \rightarrow \lambda_x x_0 \rightarrow (\lambda_x)^2 x_0 \rightarrow \cdots, \\ y: y_0 \rightarrow \lambda_y y_0 \rightarrow (\lambda_y)^2 y_0 \rightarrow \cdots, \end{cases} \quad (80)$$

where  $x_0$  and  $y_0$  are the initial values of the transformation. Let us denote  $x$  and  $y$  transformed  $n$  times by  $x_n$  and  $y_n$ , respectively.

Accordingly,  $x_n$  and  $y_n$  are defined as

$$\begin{cases} x_n := (\lambda_x)^n x_0 = x_0 e^{(\log \lambda_x)n}, \\ y_n := (\lambda_y)^n y_0 = y_0 e^{(\log \lambda_y)n}, \end{cases} \quad (81)$$

which constitute a combination of exponentials. Therefore, taking  $A = y_0$ ,  $a = \log \lambda_y$ ,  $B = x_0$ , and  $b = \log \lambda_x$  in Equation (75), we can write down  $y_n$  as a function of  $x_n$  as

$$y_n = y_0 \left( \frac{x_n}{x_0} \right)^{\frac{\log \lambda_y}{\log \lambda_x}} \propto x_n^{\frac{d_y}{d_x}}, \quad (82)$$

where the exponent of power law is  $\frac{d_y}{d_x}$ . We emphasize that this is a simple consequence of the dimensional analysis.

Furthermore, if  $y = p(x)$ , the two geometric progressions (Equation 80) lead to

$$\lambda_y y = p(\lambda_x x), \quad (83)$$

which satisfies the scale-free property (Equation 2) with  $b = \lambda_x$  and  $f(b) = \lambda_y$ . Namely, the two geometric progressions, equivalent to a combination of exponentials by the scale transformation, assures that the scale-free property holds<sup>25</sup>.

## 5. CONCLUSIONS

We have reviewed nine generating mathematical mechanisms of power laws (i.e., Yule process, Simon process, Barabási–Albert model, geometric Brownian motion with a reflecting wall and reset events, Kesten process, Generalized Lotka–Volterra model, and Bouchaud–Mézard model, and the combination of exponentials) that are widely applied in the social sciences. Since these mechanisms are only prototypes, the exponents of the power laws derived from them may not match those of real phenomena (e.g., number of links on the WWW, and so on). As explained in this paper, however, these mechanisms have been improved so that the exponents match those of real phenomena, while the basic principles of the improved mechanisms remain the same. Though many power laws as macro behaviors of systems have been studied, the mechanisms generating them from micro dynamics are not yet completely understood. In physics, however, the understanding of thermodynamics of macro behavior from quantum mechanics of micro dynamics has been advanced considerably based on statistical mechanics. A similar development may also be possible in the study of power laws.

## AUTHOR CONTRIBUTIONS

TK: Designed the overall direction of this review paper, and found out the existing models, which are highly applicable from the viewpoint of the Computational Social Science; S-IK: Surveyed existing model studies and summarized the mathematical mechanisms of those models; S-IK and TK: Wrote the manuscript.

<sup>25</sup>We thank Prof. Ken-Ichi Aoki for pointing out this observation.

## ACKNOWLEDGMENTS

The authors greatly appreciate stimulating discussions about the scale-free property in critical phenomena

with Prof. Ken-Ichi Aoki. Financial support from the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 15H05729) is gratefully acknowledged.

## REFERENCES

- Newman ME. Power laws, Pareto distributions and Zipf's law. *Contemp Phys.* (2005) **46**:323–51. doi: 10.1016/j.cities.2012.03.001
- Auerbach, F. Das Gesetz der Bevölkerungskonzentration. In: *Petermanns Geographische Mitteilungen* (1913) **59**:74–76. Available online at: [http://pubman.mpdl.mpg.de/pubman/item/escidoc:2271118/component/escidoc:2271116/Auerbach\\_1913\\_Gesetz.pdf](http://pubman.mpdl.mpg.de/pubman/item/escidoc:2271118/component/escidoc:2271116/Auerbach_1913_Gesetz.pdf)
- Estoup JB. *Gammes Sténographiques*. (1916). Paris: Institut Sténographique de France
- Zipf G. *Human Behavior the Principle of Least Effort: An Introduction to Human Ecology*. Reading, MA: Addison-Wesley (1949).
- Lotka AJ. The frequency distribution of scientific productivity. *J Wash Acad Sci.* (1926) **16**:317–23.
- Price DJDS. Networks of scientific papers. *Science* (1965) **149**:510–5.
- Willis JC. *Age and Area*. Cambridge: The University Press (1922).
- Willis JC, Yule GU. Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature* (1922) **109**:177–9. doi: 10.1038/109177a0
- Barabási AL, Albert R, Jeong H. Diameter of the world-wide web. *Nature* (1999) **401**:130–1.
- Pareto V. *Cours D'économie Politique*. Geneva: Droz (1996).
- Ijiri Y, Simon HA. *Skew Distributions and the Sizes of Business Firms*. Vol. 24. Amsterdam: North-Holland Publishing Company (1977).
- Stanley MH, Buldyrev SV, Havlin S, Mantegna RN, Salinger MA, Stanley HE. Zipf plots and the size distribution of firms. *Econ Lett.* (1995) **49**:453–7.
- Axtell RL. Zipf distribution of US firm sizes. *Science* (2001) **293**:1818–20. doi: 10.1126/science.1062081
- Gabaix X, Landier A. Why has CEO pay increased so much? *Q J Econ.* (2008) **123**:49–100. doi: 10.1162/qjec.2008.123.1.49
- Luttmer EG. Selection, growth, and the size distribution of firms. *Q J Econ.* (2007) **122**:1103–44. doi: 10.1162/qjec.122.3.1103
- Fujiwara Y. Zipf law in firms bankruptcy. *Phys A Stat Mech Appl.* (2004) **337**:219–30. doi: 10.1016/j.physa.2004.01.037
- Okuyama K, Takayasu M, Takayasu H. Zipf's law in income distribution of companies. *Phys A Stat Mech Appl.* (1999) **269**:125–31.
- Gutenberg B, Richter CF. Frequency of earthquakes in California. *Bull Seismol Soc Am.* (1944) **34**:185–8.
- Malamud BD, Morein G, Turcotte DL. Forest fires: an example of self-organized critical behavior. *Science* (1998) **281**:1840–2.
- Yule GU. A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS. *Philos Trans R Soc Lond Ser B* (1925) **213**:21–87.
- Simon HA. On a class of skew distribution functions. *Biometrika* (1955) **42**:425–40.
- Barabási AL, Albert R. Emergence of scaling in random networks. *Science* (1999) **286**:509–12.
- Gabaix X. Zipf's law for cities: an explanation. *Q J Econ.* (1999) **114**:739–67.
- Manrubia SC, Zanette DH. Stochastic multiplicative processes with reset events. *Phys Rev E* (1999) **59**:4945.
- Reed WJ. The Pareto, Zipf and other power laws. *Econ Lett.* (2001) **74**:15–19. doi: 10.1016/S0165-1765(01)00524-9
- Kesten H. Random difference equations and renewal theory for products of random matrices. *Acta Math.* (1973) **131**:207–48.
- Solomon S, Richmond P. Stable power laws in variable economies; Lotka-Volterra implies Pareto-Zipf. *Eur Phys J B* (2002) **27**:257–61. doi: 10.1140/epjb/e20020152
- Richmond P, Solomon S. Power laws are disguised Boltzmann laws. *Int J Modern Phys C* (2001) **12**:333–43. doi: 10.1142/S0129181301001754
- Malcai O, Biham O, Richmond P, Solomon S. Theoretical analysis and simulations of the generalized Lotka-Volterra model. *Phys Rev E* (2002) **66**:031102. doi: 10.1103/PhysRevE.66.031102
- Bouchaud JP, Mézard M. Wealth condensation in a simple model of economy. *Phys A Stat Mech Appl.* (2000) **282**:536–45. doi: 10.1016/S0378-4371(00)00205-3
- Miller GA. Some effects of intermittent silence. *Am J Psychol.* (1957) **70**:311–14.
- Bak P, Tang C, Wiesenfeld K. Self-organized criticality: an explanation of the 1/f noise. *Phys Rev Lett.* (1987) **59**:381.
- Carlson JM, Doyle J. Highly optimized tolerance: a mechanism for power laws in designed systems. *Phys Rev E* (1999) **60**:1412.
- Carlson JM, Doyle J. Highly optimized tolerance: robustness and design in complex systems. *Phys Rev Lett.* (2000) **84**:2529. doi: 10.1103/PhysRevLett.84.2529
- Mitzenmacher M. A brief history of generative models for power law and lognormal distributions. *Int Math.* (2004) **1**:226–51. doi: 10.1080/15427951.2004.10129088
- Sornette D. *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*. Heidelberg: Springer Science & Business Media (2006).
- Hayashi Y, Ohkubo J, Fujiwara Y, Kamibayashi N, Ono N, Yuta K, et al. *Network Kagaku No Dougubako [Tool Box of Network Science]*. Tokyo: Kindaikagakusya (2007).
- Farmer JD, Geanakoplos J. *Power Laws in Economics and Elsewhere*. Technical Report, Santa Fe Institute (2008).
- Gabaix X. Power laws in economics and finance. *Annu Rev Econ.* (2009) **1**:255–94. doi: 10.3386/w14299
- Gabaix X. Power laws in economics: an introduction. *J Econ Perspect.* (2016) **30**:185–205. doi: 10.1257/jep.30.1.185
- Simkin MV, Roychowdhury VP. Re-inventing willis. *Phys Rep.* (2011) **502**:1–35. doi: 10.1016/j.physrep.2010.12.004
- Pinto CM, Lopes AM, Machado JT. A review of power laws in real life phenomena. *Commun Nonlinear Sci Numer Simul.* (2012) **17**:3558–78. doi: 10.1016/j.cnsns.2012.01.013
- Piantadosi ST. Zipf's word frequency law in natural language: a critical review and future directions. *Psychon Bull Rev.* (2014) **21**:1112–30. doi: 10.3758/s13423-014-0585-6
- Machado JT, Pinto CM, Lopes AM. A review on the characterization of signals and systems by power law distributions. *Signal Process.* (2015) **107**:246–53. doi: 10.1016/j.sigpro.2014.03.003
- Slanina F. *Essentials of Econophysics Modelling*. Oxford: Oxford University Press (2013).
- Gibrat R. *Les Inégalités Économiques*. Paris: Recueil Sirey (1931).
- Merton RK. The Matthew effect in science. *Science* (1968) **159**:56–63.
- Price DdS. A general theory of bibliometric and other cumulative advantage processes. *J Assoc Inform Sci Technol.* (1976) **27**:292–306.
- Bacaré N. Yule and evolution (1924). In: *A Short History of Mathematical Population Dynamics*. London: Springer (2011). p. 81–8. Available online at: <http://www.springer.com/gb/book/9780857291141>
- Kimmel M, Axelrod DE. *Branching Processes in Biology*. New York, NY: Springer Publishing Company, Incorporated (2016).
- Risken H. Fokker-planck equation. In: *The Fokker-Planck Equation*. Berlin: Springer (1996). p. 63–95. Available online at: <http://www.springer.com/la/book/9783540615309>
- Richmond P. Power law distributions and dynamic behaviour of stock markets. *Eur Phys J B* (2001) **20**:523–26. doi: 10.1007/PL00011108
- Yakovenko VM, Rosser Jr JB. Colloquium: statistical mechanics of money, wealth, and income. *Rev Mod Phys.* (2009) **81**:1703. doi: 10.1103/RevModPhys.81.1703
- Banerjee A, Yakovenko VM. Universal patterns of inequality. *N J Phys.* (2010) **12**:075032. doi: 10.1088/1367-2630/12/7/075032
- Toda AA. *Zipf's Law: A Microfoundation* (2017). Available online at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2808237](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2808237)



56. Reed WJ. The Pareto law of incomes—an explanation and an extension. *Phys A Stat Mech Appl.* (2003) **319**:469–86. doi: 10.1016/S0378-4371(02)01507-8
57. Marsili M, Maslov S, Zhang YC. Dynamical optimization theory of a diversified portfolio. *Phys A Stat Mech Appl.* (1998) **253**:403–18.
58. Solomon S. Stochastic lotka-volterra systems of competing auto-catalytic agents lead generically to truncated pareto power wealth distribution, truncated levy distribution of market returns, clustered volatility, booms and crashes. *arXiv preprint cond-mat/9803367* (1998).
59. Malcai O, Biham O, Solomon S. Power-law distributions and Levy-stable intermittent fluctuations in stochastic systems of many autocatalytic elements. *Phys Rev E* (1999) **60**:1299.
60. Wilson KG. Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. *Phys Rev B* (1971) **4**:3174.
61. Pinsky M, Karlin S. *An Introduction to Stochastic Modeling*. Cambridge, MA: Academic Press (2010).
62. Osaki S. *Applied Stochastic System Modeling*. Heidelberg: Springer Science & Business Media (1992).

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kumamoto and Kamiyigashi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



## APPENDIX

Some mathematical supplements are given in this appendix.

### A.1. Poisson Process

We consider the Poisson process [61, 62] with the Poisson rate  $\lambda$  (a positive constant), that is, the events occur on average  $\lambda$  times per unit time. The probability that an event occurs  $n$  times in  $(t, t+h]$  follows the Poisson distribution:

$$P\{\hat{N}(t+h) - \hat{N}(t) = n; \text{ rate } \lambda\} = e^{-\lambda h} \frac{(\lambda h)^n}{n!}, \quad (n = 0, 1, 2, \dots) \quad (\text{A.1})$$

where  $\hat{N}(t)$  denotes the number of times that the events occur in  $(0, t]$ . When  $h$  is the infinitesimal time interval, the probabilities of event occurrence can be expressed by  $o(h^k)$ . The probability that no event occurs in  $(t, t+h]$  is

$$\begin{aligned} P\{\hat{N}(t+h) - \hat{N}(t) = 0; \text{ rate } \lambda\} &= e^{-\lambda h} \\ &= 1 - \lambda h + o(h). \left( \lim_{h \rightarrow 0} \frac{o(h^k)}{h^k} = 0 \right) \end{aligned} \quad (\text{A.2})$$

Similarly, the probability that one event occurs is

$$\begin{aligned} P\{\hat{N}(t+h) - \hat{N}(t) = 1; \text{ rate } \lambda\} &= e^{-\lambda h} \lambda h = (1 - \lambda h + o(h)) \lambda \\ h &= \lambda h + o(h), \end{aligned} \quad (\text{A.3})$$

and the probability that more than two events occur is

$$\begin{aligned} P\{\hat{N}(t+h) - \hat{N}(t) \geq 2; \text{ rate } \lambda\} &= \sum_{n=2}^{\infty} e^{-\lambda h} \frac{(\lambda h)^n}{n!} \\ &= e^{-\lambda h} \left( \sum_{n=0}^{\infty} \frac{(\lambda h)^n}{n!} - 1 - \lambda h \right) \\ &= 1 - e^{-\lambda h} (1 + \lambda h) \\ &= 1 - (1 - \lambda h + o(h)) (1 + \lambda h) \\ &= o(h). \end{aligned} \quad (\text{A.4})$$

### A.2. Pure Birth Process

We generalize the Poisson process so that the Poisson rate depends on the number of times that the events have already occurred. To apply this generalized Poisson process to the evolution model in biology, we interpret the occurrence of events as the births of new species without deaths [61, 62].

First, we are interested in the probability that the number of species becomes  $n$  ( $\in \mathbb{N}$ ) at time  $t$  ( $\in \mathbb{R}$ ) with the initial number of species  $n_{s0}$  at time 0. It is denoted by  $p(n, t) \left( := P\{\hat{N}_s(t) = n\} \right)$ , where  $\hat{N}_s(t)$  is the number of species at time  $t$ . Then, we derive the time evolution equation of  $p(n, t)$ . The probability  $p(n, t+h)$  is given as the sum of the following probabilities:

- the probability that  $\hat{N}_s(t) = n$  and no birth occurs in  $(t, t+h]$  with rate  $\Lambda_s(n)$ ;

- the probability that  $\hat{N}_s(t) = n-1$  and one birth occurs in  $(t, t+h]$  with rate<sup>26</sup>  $\Lambda_s(n-1)$ ;
- the probability that  $\hat{N}_s(t) = n-2$  and two births occur in<sup>27</sup>  $(t, t+h]$ ;
- the probability that  $\hat{N}_s(t) = n-k$  and  $k$  births occur in  $(t, t+h]$ ;

where  $\Lambda_s(n)$  is the Poisson rate when the number of species is  $n$ . Accordingly, we obtain

$$\begin{aligned} p(n, t+h) &= P\{\hat{N}_s(t) = n \cap \text{no birth occurs in } (t, t+h] \\ &\quad \text{with } \Lambda_s(n)\} \\ &\quad + P\{\hat{N}_s(t) = n-1 \cap 1 \text{ birth occurs in } (t, t+h] \\ &\quad \text{with } \Lambda_s(n-1)\} \\ &\quad + \sum_{k=2}^{n-n_{s0}} P\{\hat{N}_s(t) = n-k \cap k \text{ births occur in } \\ &\quad (t, t+h]\}. \end{aligned} \quad (\text{A.5})$$

From equations (A.2), (A.3), and (A.4), the probabilities on the right-hand side of (A.5) are expressed respectively as the orders of  $h$ :

$$\begin{cases} P\{\hat{N}_s(t) = n \cap \text{no birth occurs in } (t, t+h]\} \\ \quad = p(n, t) \times \{1 - \Lambda_s(n)h + o(h)\}, \\ P\{\hat{N}_s(t) = n-1 \cap 1 \text{ birth occurs in } (t, t+h]\} \\ \quad = p(n-1, t) \times \{\Lambda_s(n-1)h + o(h)\}, \\ \sum_{k=2}^{n-n_{s0}} P\{\hat{N}_s(t) = n-k \cap k \text{ births occur in } (t, t+h]\} = o(h). \end{cases} \quad (\text{A.6})$$

We combine (A.5) with (A.6), and obtain the difference equation:

$$\frac{p(n, t+h) - p(n, t)}{h} = -\Lambda_s(n)p(n, t) + \Lambda_s(n-1)p(n-1, t) + \frac{o(h)}{h}. \quad (\text{A.7})$$

<sup>26</sup>Strictly speaking, the Poisson rate is not constant at  $\Lambda_s(n-1)$  in  $(t, t+h]$ , that is, the Poisson rate change into  $\Lambda_s(n)$  from  $\Lambda_s(n-1)$  when the birth occurs. Therefore, the accurate probability is  $P\{\hat{N}_s(t) = n-1\} \times P\{\text{one species is born in } (t, t+j] \text{ with rate } \Lambda_s(n-1)\} \times P\{\text{no species is born in } (t+j, t+h] \text{ with rate } \Lambda_s(n)\}$ , where  $t+j$  ( $0 < j < h$ ) is the time of the birth. However, since we take the limit  $h \rightarrow 0$  at the end, even if we deal with the probability this strictly, the time evolution equation of the final result will be the same.

<sup>27</sup>Even if we precisely consider the changing Poisson rate with births, this probability will eventually be  $o(h)$ . Therefore, we do not need the precise values for the exact Poisson rate and the probabilities that  $k(\geq 2)$  species are born in  $(t, t+h]$ .

We take the limit  $h \rightarrow 0$  and obtain the ODEs with the initial conditions:

$$\begin{aligned} \text{for } n > n_{s0}, \quad & \begin{cases} \frac{\partial p(n, t)}{\partial t} = -\Lambda_s(n)p(n, t) + \Lambda_s(n-1)p(n-1, t), \\ p(n, 0) = 0, \end{cases} \\ \text{for } n = n_{s0}, \quad & \begin{cases} \frac{\partial p(n_{s0}, t)}{\partial t} = -\Lambda_s(n_{s0})p(n_{s0}, t), \\ p(n_{s0}, 0) = 1, \end{cases} \end{aligned} \quad (\text{A.8})$$

which are called the Kolmogorov's forward equations. The ODEs (A.8) can be solved and yield

$$\begin{cases} p(n, t) = \int_0^t e^{-\Lambda_s(n)(t-s)} \Lambda_s(n-1)p(n-1, s) ds & \text{for } n > n_{s0}, \\ p(n_{s0}, t) = e^{-\Lambda_s(n_{s0})t}. \end{cases} \quad (\text{A.9})$$

### A.3. Linear Birth Process

Next, we consider the linear birth process [61, 62] that is mathematically defined as a special case of the pure birth process. When the Poisson rate  $\Lambda_s(n)$  is proportional to the number of species  $n$ ,

$$\Lambda_s(n) = \lambda_s n, \quad (\text{A.10})$$

where  $\lambda_s$  is a positive constant, this pure birth process is called the linear birth process<sup>28</sup>. Then, we can interpret the birth of new species in this process as the occurrence of branching in the evolutionary tree (Figure 2). In particular, the linear birth process means that the branchings occur independently on each line of a species as the Poisson processes with the Poisson rate  $\lambda_s$ , which is common for all existing species.

The solutions of (A.9) for the Yule process can be recursively calculated and yield

$$\begin{cases} p(n, t) = \binom{n-1}{n-n_{s0}} (e^{-\lambda_s t})^{n_{s0}} (1 - e^{-\lambda_s t})^{n-n_{s0}} \\ \quad \left( \binom{n}{m} := \frac{n!}{m!(n-m)!} \right) & \text{for } n > n_{s0}, \\ p(n_{s0}, t) = e^{-\lambda_s n t}. \end{cases} \quad (\text{A.11})$$

Then, the expectation value and the variance of the number of species at time  $t$  is given by

$$\begin{aligned} E[\hat{N}_s(t)] &= \sum_{n=n_{s0}}^{\infty} n p(n, t) = n_{s0} e^{\lambda_s t}, \\ \text{Var}[\hat{N}_s(t)] &= E[\hat{N}_s(t)^2] - E[\hat{N}_s(t)]^2 = n_{s0} e^{\lambda_s t} (e^{\lambda_s t} - 1). \end{aligned} \quad (\text{A.12})$$

<sup>28</sup>Though this process is also called the Yule–Furry process, we call it the linear birth process in this paper to distinguish it from the Yule process that generates a power law.

Let  $P_s\{0 < \text{age} \leq t \text{ at } \tau\}$  be the probability of the species whose age, that is, the time intervals elapsed since the birth, is  $t$  or less at time  $\tau$  ( $> t$ ). This probability is given by

$$\begin{aligned} P_s\{0 < \text{age} \leq t \text{ at } \tau\} &= E \left[ \frac{\hat{N}_s(\tau) - \hat{N}_s(\tau - t)}{\hat{N}_s(\tau)} \right] \\ &= 1 - E \left[ \frac{\hat{N}_s(\tau - t)}{\hat{N}_s(\tau)} \right] \\ &\simeq 1 - \frac{E[\hat{N}_s(\tau - t)]}{E[\hat{N}_s(\tau)]} = 1 - e^{-\lambda_s t}, \end{aligned} \quad (\text{A.13})$$

where the approximately equal symbol holds only for a large time<sup>29</sup>  $\tau$ . Therefore, it no longer depends on  $\tau$ . Let us use  $\ell_s(s)$  to denote the probability density function for the age  $s$  of species at a large time. By the probability of the species whose age is  $t$  or less at a large time, it is defined as

$$\int_0^t \ell_s(s) ds = P_s\{0 < \text{age} \leq t \text{ at a large time}\}. \quad (\text{A.14})$$

Differentiating both sides of (A.14) with respect to  $t$ , we obtain

$$\ell_s(t) = \frac{dP_s\{0 < \text{Age} \leq t \text{ at a large time}\}}{dt} = \lambda_s e^{-\lambda_s t}. \quad (\text{A.15})$$

### A.4. Multiplicative Process

The multiplicative process is the discrete-time stochastic process defined as

$$\hat{X}(t+1) = \hat{r}(t)\hat{X}(t) \quad (t = 0, 1, 2, \dots), \quad (\text{A.16})$$

where  $\hat{r}(t)$ , for all times  $t$ , are independent and equally-distributed random variables with  $\nu := E[\log \hat{r}(t)]$  and  $\sigma^2 := \text{Var}[\log \hat{r}(t)]$ . This process is essentially equivalent to the GBM because both probability density functions are identically the log-normal distributions in the large time limit.

We can easily obtain the solution of (A.16) in the logarithmic form as follows:

$$\log \hat{X}(t) = \sum_{i=0}^{t-1} \log \hat{r}(i) + \log x_0, \quad (\text{A.17})$$

where  $x_0$  is the initial value of  $\hat{X}(t)$ . We then define the new variable  $\hat{Y}(t)$  as

$$\hat{Y}(t) := \frac{\log \hat{X}(t) - \log x_0 - t\nu}{\sqrt{t}} = \frac{\sum_{i=0}^{t-1} (\log \hat{r}(i) - \nu)}{\sqrt{t}}. \quad (\text{A.18})$$

<sup>29</sup>We consider only the probability in a large time, because we are interested in only the power-law distribution as the stationary state at a large time.

By the central limit theorem, we obtain the probability density function of  $\hat{Y}(t)$  in the time limit  $t \rightarrow \infty$ :

$$q(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{y^2}{2\sigma^2}\right], \quad (\text{A.19})$$

which is the normal distribution. Consequently, by a change of variables, we obtain the probability density function of  $\hat{X}(t)$  as follows:

$$p(x) = \frac{1}{x\sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{\{\log x - \log x_0 - t\nu\}^2}{2\sigma^2 t}\right], \quad (\text{A.20})$$

which is the same as the log-normal distribution as (28) of the GBM with  $\nu = \mu - \frac{\sigma^2}{2}$ .

### A.5. Stationary Solution of the Fokker-Planck Equation With Reflecting Wall

Here we provide a stationary solution of the FPE with reflecting wall [23, 39, 55].

The SDE<sup>30</sup> of an Itô process for the random variable  $\hat{X}(t)$  is given by

$$d\hat{X}(t) = a(\hat{X}(t), t)dt + b(\hat{X}(t), t)d\hat{B}(t), \quad (\text{A.21})$$

where  $\hat{B}(t)$  is a standard Brownian motion;  $E[d\hat{B}(t)] = 0$ ,  $\text{Var}[d\hat{B}(t)] = dt$ . This SDE is equivalent to the Langevin equation [51]:

$$\frac{d\hat{X}(t)}{dt} = a(\hat{X}(t), t) + b(\hat{X}(t), t)\hat{\Gamma}(t), \quad (\text{A.22})$$

where the noise term  $\hat{\Gamma}(t)$  satisfies

$$\begin{cases} E[\hat{\Gamma}(t)] = 0, \\ E[\hat{\Gamma}(t)\hat{\Gamma}(t')] = \delta(t - t'). \end{cases} \quad (\text{A.23})$$

We can obtain the FPE for the random variable  $\hat{X}(t)$  with the probability density  $p(x, t)$  as

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x}\{a(x, t)p(x, t)\} + \frac{\partial^2}{\partial x^2}\left\{\frac{b(x, t)^2}{2}p(x, t)\right\}. \quad (\text{A.24})$$

Then we define the flux  $J(x, t)$  as

$$J(x, t) := a(x, t)p(x, t) - \frac{\partial}{\partial x}\left\{\frac{b(x, t)^2}{2}p(x, t)\right\}, \quad (\text{A.25})$$

so that we can interpret (A.24) as the continuity equation

$$\frac{\partial p(x, t)}{\partial t} + \frac{\partial J(x, t)}{\partial x} = 0. \quad (\text{A.26})$$

<sup>30</sup>In the Stratonovich convention this SDE is represented by  $d\hat{X}(t) = \left\{a(\hat{X}(t), t) - \frac{1}{2}b(\hat{X}(t), t)\frac{\partial b(\hat{X}(t), t)}{\partial \hat{X}(t)}\right\}dt + b(\hat{X}(t), t) \circ d\hat{B}(t)$ .

When  $a(x, t)$  and  $b(x, t)$  are the time-independent functions, that is,  $a(x, t) = a(x)$  and  $b(x, t) = b(x)$ , the stationary solution  $p(x)$  is defined by the condition<sup>31</sup>

$$\frac{\partial p(x)}{\partial t} = 0, \quad (\text{A.27})$$

that is equivalent to

$$\frac{\partial J(x)}{\partial x} = 0, \quad (\text{A.28})$$

where  $J(x)$  is the stationary flux. Accordingly, the stationary flux  $J(x)$  must be constant.

When the stationary flux  $J(x)$  takes a nonzero value, the stationary state means that particles flow in from one side of infinity and out the other side. This situation causes the stationary probability density function  $p(x)$  to be nonzero at  $x = \pm\infty$ . Consequently, the nonzero stationary flux cannot give us a power-law probability density function that can be normalized, because any power function blows up at one side of infinity. In contrast, when the stationary flux  $J(x)$  vanishes anywhere, we can set the reflecting wall at  $x = x_{\min}$  so that the stationary probability density function  $p(x)$  vanishes outside of the wall. The reflecting wall enables us to obtain a power-law probability density function that can be normalized, because we can cut out the side of infinity where the power function blows up. For this reason, we consider only the case that the flux vanishes at a boundary, that is, the reflecting wall.

In this case, we obtain the second-order ODE

$$J(x) = a(x)p(x) - \frac{d}{dx}\left\{\frac{b(x)^2}{2}p(x)\right\} = 0 \quad (\text{A.29})$$

that the stationary solution  $p(x)$  satisfies.

The stationary solution is obtained as the solution of (A.29):

$$\begin{cases} p(x) = p(x_0)b(x_0)^2e^{f(x)}, \\ f(x) := -2\log\{b(x)\} + \int_{x_0}^x \frac{2a(x')}{b(x')^2}dx', \end{cases} \quad (\text{A.30})$$

where  $x_0 (\geq x_{\min})$  is an arbitrary constant. If  $a(x)$  and  $b(x)$  are the power functions that satisfy the condition

$$\frac{a(x)}{b(x)^2} \propto \frac{1}{x}, \quad (\text{A.31})$$

namely,

$$\begin{cases} a(x) = ax^{2n-1} \quad (a: \text{constant}), \\ b(x) = bx^n \quad (b: \text{constant}), \end{cases} \quad (\text{A.32})$$

we obtain the stationary solution as the power function of  $x$ :

$$p(x) = Cx^{-\alpha} \quad \left(C := p(x_0)x_0^\alpha, \quad \alpha := 2n - \frac{2a}{b^2}\right). \quad (\text{A.33})$$

<sup>31</sup>Though the existence of the stationary solution is nontrivial, we assume it here.

This stationary solution  $p(x)$  must satisfy the normalization condition

$$1 = \int_{x_{\min}}^{\infty} p(x) dx, \quad (\text{A.34})$$

where we set the reflecting wall at  $x = x_{\min} (> 0)$  and assume  $\alpha > 1$ . The normalization condition

$$1 = \int_{x_{\min}}^{\infty} p(x) dx = \frac{C}{\alpha - 1} (x_{\min})^{-\alpha+1}. \quad (\text{A.35})$$

determines the constant  $C$  as

$$C = (\alpha - 1)(x_{\min})^{-\alpha+1}. \quad (\text{A.36})$$

Thus, we have the stationary solution

$$p(x) = (\alpha - 1)(x_{\min})^{-\alpha+1} x^{-\alpha} \quad \left( \alpha = 2n - \frac{2a}{b^2} > 1 \right). \quad (\text{A.37})$$