Poisson distributions in stochastic dynamics of gene expression: What events do they count?

Julian Lee*

Department of Bioinformatics and Life Science, Soongsil University, Seoul, Korea

(Dated: August 16, 2023)

Abstract

The Poisson distribution is the probability distribution of the number of independent events in a given period. Although the Poisson distribution appears ubiquitously in various stochastic dynamics of gene expression, both as time-dependent distributions and stationary distributions, underlying independent events that give rise to such distributions have not been clear, especially in the presence of the degradation of gene products, which is not a Poisson process. I show that the variable following the Poisson distribution is the number of independent events where biomolecules are created, destined to survive until the end of a given time duration. This new viewpoint enables me to rederive the Poisson distribution as a time-dependent probability distribution for molecule numbers in various monomolecular reaction models of stochastic gene dynamics. Additionally, it allows me to derive an analytic form of the time-dependent probability distribution for multispecies monomolecular reaction models with species whose lifetimes follow nonexponential distributions, which is the convolution of the Poisson distribution with the multinomial distribution. This distribution is then utilized for deriving a novel series expansion form of a time-dependent distribution for a model with a stochastic production rate.

Keywords: Gene Regulatory Network, Stochastic fluctuation, Poisson distribution

* jul@ssu.ac.kr

I. INTRODUCTION

Gene regulatory network(GRN) controls the life process by producing and degrading various kinds of proteins that perform important biological functions. It is a well-known fact that the time evolution of mRNA and/or protein molecules in such a network is stochastic [1–27]. Even when various external conditions such as cellular environments are identical, there is always intrinsic noise due to remaining uncontrolled factors that influence the GRN of interest, making its dynamics stochastic. It has been suggested that biological organisms may have evolved to take advantage of such fluctuations [13].

In simple theoretical models of stochastic GRN dynamics, the Poisson distribution,

$$P_{\text{Poisson}}(n;\mu) \equiv e^{-\mu}\frac{\mu^n}{n!}, \tag{1}$$

often appears as a probability distribution for the number of mRNA or protein molecules, both as time-dependent and stationary distributions [1, 3, 16, 17, 23, 25, 26, 28–30]. In fact, the Poisson distribution arises from a Poisson process, where the probability of an event occurring during a short time interval $[t, t+dt]$ is independent of events in other time regions [31–33]. Now, consider a simple transcription process where a gene is always active, and an mRNA molecule X is transcribed with a rate α,

$$\varnothing \xrightarrow{\alpha} X. \tag{2}$$

The creation events are indeed independent, forming a Poisson process. Therefore, if we start from zero molecules at $t = 0$, then the number n of mRNA molecules at any later time t is the same as the number of creation events during the time interval $[0, t]$, and consequently, it follows the Poisson distribution. However, consider a model where a degradation

$$X \xrightarrow{\beta} \varnothing. \tag{3}$$

is also included. Now, the degradation is *not* a Poisson process, although sometimes it is erroneously described as such in the literature. Since a molecule that already got degraded cannot be degraded again, the probability of a degradation event happening during the short period $[t, t+dt]$ depends on how many degradation events happened in times earlier than t. However, the Poisson distributions still appear ubiquitously in this class of models as the distributions of the number of mRNA or protein molecules [1, 3, 16, 17, 23, 25, 26, 28–30]. Furthermore, it has been also shown that an arbitrary time-dependent probability

distribution of molecule numbers in a monomolecular reaction network is a superposition of convolutions of a Poisson distribution with a multinomial distribution [29]. The universality of the Poisson distribution in stochastic gene dynamics has also been explained in terms of queuing theory and non-linear transformation of time [30], but no explicit connection has been made between the number of molecules and the number of some independent events.

Therefore, I address the following question in this paper: Given that the Poisson distribution appears so ubiquitously in the stochastic dynamics of GRN as the distribution of molecule number of mRNA or protein, is this molecule number equal to the number of certain independent events that happened during a given time interval? As I will show, the answer to this question is affirmative. In fact, it is the number of events where mRNA molecules are created, which are destined to survive until the end of a given interval. The answer is very simple once stated, almost on the verge of being trivial. The most probable reason that it has rarely been discussed in the literature is that it may have been counter-intuitive to consider a birth of a particle with a given fate. However, it is important to note that although the fate of a molecule is not determined at the time of its creation, the *probability* of its given fate at the end of the time interval, death or survival, is already determined at the time of its creation, and it is all that matters in defining a Poisson process.

Using this new viewpoint, I can not only rederive Poisson distributions as molecule number distributions for models of gene expression with time-dependent rates and a model with time-delayed degradation, but also derive an analytic form of the molecule number distribution for multispecies monomolecular reaction model that includes species with non-exponential distributions of lifetimes, which is a convolution of the Poisson distribution with the multinomial distribution. Furthermore, by taking the weighted average of these distributions, I can also derive a novel series expansion form of the molecule number distribution for a model with a stochastic production rate.

The remainder of the paper is organized according to the order of increasing complexity. In section II, I will briefly review Poisson processes, where I will emphasize the importance of inhomogeneous Poisson processes, consisting of independent events that do not necessarily follow identical distributions. In section III, I will consider molecule creations without degradation, a textbook example of a Poisson process. The shifted Poisson distribution will also be introduced for a nonzero number of initial molecules. In section IV, I will describe the molecule degradation, which is *not* a Poisson process. I show that a general

time-dependent distribution of molecule numbers can be expressed as a superposition of binomial distributions. In section V, I will describe molecule creations with degradations and the underlying Poisson process. I show that any time-dependent distribution can be expressed in terms of the shifted Poisson and the binomial distributions. In section VI, a non-Markovian model with delayed degradation will be considered, where I will not only rederive the Poisson distribution derived in earlier work with less effort [17] but also derive a more general time-dependent distribution by convoluting the Poisson distribution with the multinomial distribution. In section VII, a multispecies monomolecular reaction network model that includes species with finite upper limits on their lifetimes is studied, and derive a time-dependent distribution of molecule numbers in the form of a convolution of a Poisson distribution with a multinomial distribution. This model is generalized in section VIII to a multispecies monomolecular reaction network model that includes species with non-exponential distributions of lifetimes. This model encompasses all the models in the previous sections as special cases, and again I find a general time-dependent distribution of molecule numbers in the form of a convolution of a Poisson distribution with a multinomial distribution. A model with a stochastic production rate, the one-species telegraph model, will be considered in section IX, where the probability distribution of the molecule number is neither Poisson nor its convolution with a binomial distribution. However, by taking linear combinations of the distributions derived in the previous sections with different parameter values, it will be shown that for a model without degradation, a time-dependent distribution for *arbitrary* initial condition can be expressed in terms of convolutions of confluent hypergeometric functions and binomial distributions, which leads to a novel series representation of the distribution. Finally, the discussion is given in section IX.

II. THE POISSON DISTRIBUTION AND POISSON PROCESSES

The Poisson distribution in Eq.(1) is specified by a single parameter μ, the expected number of independent events in a given time duration. If we partition the time interval $[0, t]$ into N sub-intervals of small size $\Delta t \equiv t/N$, then the probability of events happening more than once in each sub-interval is $O\left((\Delta t)^2\right)$, and the probability of a single occurrence of the event takes the form $p = \lambda \Delta t + O\left((\Delta t)^2\right)$ when such a probability is time-independent. Therefore,

the total number of events in $[0,t]$ approximately follows the binomial distribution[1],

$$P_{\rm binom}(n;\{N,p\}) \equiv \frac{N!}{n!(N-n)!}p^n(1-p)^{N-n}, \tag{4}$$

the distribution of the number of successes in N independent and identical trials, with a success probability of p at each trial. The Poisson distribution is recovered by taking the limit of $N\to\infty$ with $\mu(t)=Np=\lambda t$ fixed (Appendix A):

$$P_{\rm Poisson}(n;\mu(t)) \equiv \frac{e^{-\mu(t)}}{n!}\mu(t)^n, \tag{5}$$

where $\mu(t)$ is the expected number of events in the time-interval $[0,t]$.

An important point to note here is that the condition of identical trials is not required to obtain the Poisson distribution. We only have to require *independent* trials [31, 32]. In this case, the binomial distribution in Eq.(4) is generalized to

$$\tilde{P}_{\rm\ binom}(n;\{p_1,\cdots,p_N\}) \equiv \sum_{\{i_1<i_2\cdots<i_n\}} p_{i_1}p_{i_2}\cdots p_{i_n}\prod_{k\notin\{i_1,\cdots,i_n\}}(1-p_k) \tag{6}$$

where the probability of success at i-th trial is p_i, and the summation is over all distinct set of n indices $\{i_1<i_2,\cdots<i_n\}$. Again, the distribution in Eq.(6) becomes the Poisson distribution in the limit of $\Delta t\to 0$, now with the time-dependent function $\lambda(t)$, so that $\mu(t)=\int_0^t\lambda(t')dt'$ [32](Appendix A). This is the process where the probability of an event happening in an infinitesimal time interval $[t,t+dt]$ is $\lambda(t)dt$, called an inhomogeneous Poisson process, to distinguish it from the case with constant λ, called a homogeneous Poisson process.

III. MOLECULE CREATION WITHOUT DEGRADATION

This is a simple birth process

$$\varnothing \xrightarrow{\alpha(t)} X, \tag{7}$$

[1] The range of n in the probability distribution $P(n,t)$ will be considered to be all the integers without any restriction in this work unless specified otherwise. This is acceptable as long as we set $P(n,t)=0$ for illegitimate values of n. All the explicit forms of probability distributions considered here such as binomial, multinomial, and Poisson distributions contain factorials of negative integers in the denominators whenever the value of the molecule number is out of legitimate range, and vanish because $(j!)^{-1}=[\Gamma(j+1)]^{-1}=0$ whenever j is a negative integer. Taking the range of the molecule number to be the whole integer is especially convenient for summation, where the summation index can be shifted freely, which I will do throughout the text without further elaboration.

where X describes an mRNA or protein molecule, and the creation rate $\alpha(t)$ is time-dependent in general. Since molecule creations at distinct time points are independent of each other, these creation events form a Poisson process. If the number of molecules is zero at $t = 0$, then the molecule number n at a later time $t > 0$ is equal to the number of creation events in the time interval $[0, t]$. Therefore, the probability distribution $P(n, t|0)$ of n, under the condition of vanishing initial molecule number, is given by the Poisson distribution $P(n, t|0) = P_{\rm Poisson}(n; \mu(t))$, with $\mu(t) = \int_0^t \alpha(t')dt'$.

Even when the number of molecules takes a non-zero value n_0 at $t = 0$, the number of molecules n at a later time $t > 0$ is simply $n_0 + n'$ where n' is the number of creation events in the time interval $[0, t]$. Therefore, $n' = n - n_0$ still follows the Poisson distribution, leading to the distribution for n given by

$$P(n, t|n_0) = \frac{e^{-\mu(t)}}{(n - n_0)!}\mu(t)^{n-n_0}. \tag{8}$$

Eq.(8) is what I will call a shifted Poisson distribution.

Up to now, we wrote down the specific probability distribution for the model, but one can also consider the master equation describing the probability distribution dynamics,

$$\frac{\partial P(n, t)}{\partial t} = \alpha(t)\left[P(n - 1, t) - P(n, t)\right]. \tag{9}$$

It is straightforward to check that Eq.(8) is a solution of the master equation (9) by direct substitution. By the linearity of the master equation (9), an analytic form of the general solution with an arbitrary initial distribution $P(n, 0) = v(n)$ can be constructed by superposing the expressions in Eq.(8),

$$P(n, t) = \sum_{n_0} P(n, t|n_0)v(n_0) = \sum_{n_0} \frac{e^{-\mu(t)}}{(n - n_0)!}\mu(t)^{n-n_0}v(n_0). \tag{10}$$

As we will see, arbitrary time-dependent distributions in various models of stochastic gene dynamics can be naturally expanded in terms of shifted Poisson distribution, similar in the form to Eq.(10).

From Eq.(10), we see that the master equation (9) admits stationary solutions

$$P_{\rm st}(n) = \sum_{n_0} \frac{e^{-\mu(\infty)}}{(n - n_0)!}\mu(\infty)^{n-n_0}v(n_0) \tag{11}$$

that depend on initial distributions, if and only if $\mu(\infty) = \int_0^\infty \alpha(s)ds$ exists. In particular, for a constant rate α, there is no stationary solution because $\mu(t) = \alpha t$ increases indefinitely

with time, and $P(n,t)$ converges to zero pointwise for all n. In this case, all the states in the Markov chain are transient states [34].

IV. MODEL ONLY WITH DEGRADATION

This is a simple death process

$$X \xrightarrow{\beta(t)} \varnothing. \tag{12}$$

In contrast to the creation process in the previous section, this process is *not* a Poisson process, because any molecule that has been degraded cannot be degraded again. Therefore, the degradation events are not independent, and the probability of a degradation event depends on the number of molecules available. The master equation describing the dynamics of this model is

$$\frac{\partial P(n,t)}{\partial t} = \beta(t)\left[(n+1)P(n+1,t) - nP(n,t)\right]. \tag{13}$$

To get a non-zero number of molecules, the initial number n_0 of molecules must be non-zero. The number of molecules at later times can be obtained by using the fact that the event of a given molecule surviving at a later time t is independent of what happens to other molecules, whose survival probability is given by

$$p_{\text{surv}}(t|0) = \exp\left(-\int_0^t \beta(s)ds\right). \tag{14}$$

Therefore, the number n of surviving molecules at time t for given value of n_0 follows the binomial distribution

$$P(n,t|n_0) = P_{\text{binom}}(n;\{n_0, p_{\text{surv}}(t|0)\}) \equiv \frac{n_0!}{n!(n_0-n)!} p_{\text{surv}}(t|0)^n (1-p_{\text{surv}}(t|0))^{n_0-n}. \tag{15}$$

It can be checked that the expression in Eq.(15) is the solution of the master equation (13) by direct substitution. Again, the general solution with an arbitrary initial distribution $P(n,0) = v(n)$ can be constructed by the superposition

$$P(n,t) = \sum_{n_0} P(n,t|n_0)v(n_0) = \sum_{n_0} \frac{n_0!}{n!(n_0-n)!} p_{\text{surv}}(t|0)^n (1-p_{\text{surv}}(t|0))^{n_0-n} v(n_0). \tag{16}$$

A Poisson distribution appears in the special case where the initial distribution is Poissonian:

$$v(n) = \frac{e^{-\mu_0}\mu_0^n}{n!}, \tag{17}$$

in which case we obtain the Poisson distribution $P(n,t) = P_{\rm Poisson}(n;\mu(t))$ with

$$\mu(t) = \mu_0 p_{\rm surv}(t|0), \tag{18}$$

because the superposition of binomial distributions weighted by a Poisson distribution is again a Poisson distribution [32](Appendix B). In the next section, I will show that this distribution can be interpreted as a result of a Poisson process.

When the integral $\int_0^t \beta(s)ds$ diverges as $t \to \infty$, as in the case of a constant β, we get $\lim_{t\to\infty} p_{\rm surv}(t|0) = 0$. In this case, we see from Eq.(16) that the probability distribution approaches the stationary distribution

$$\lim_{t\to\infty} P(n,t) = P_{\rm st}(n) = \delta_{n,0}, \tag{19}$$

regardless of the initial distribution.

V. MODEL WITH BOTH CREATION AND DEGRADATION

This model is described by

$$\varnothing \xrightarrow{\alpha(t)} X, \quad X \xrightarrow{\beta(t)} \varnothing, \tag{20}$$

which belongs to a class of models called the birth-death models [31–33]. Let us first consider the case where the number of X molecules is initially zero. Molecules are created in the time interval $[0,t]$, but only a fraction of them survive at t, which I will call surviving molecules. When a molecule is created in the short time interval $[t', t'+dt']$ with $t' < t$, its fate at t is undetermined, but the probability of its survival $p_{\rm surv}(t|t')$ at time t is already determined to be

$$p_{\rm surv}(t|t') = \exp\left(-\int_{t'}^{t} \beta(s)ds\right), \tag{21}$$

and consequently, the probability that a surviving molecule is created in $[t', t'+dt']$ is given by $p_{\rm surv}(t|t')\alpha(t')dt'$, independent of events that happen in other regions of time. Therefore, the number of molecules at a later time t is equal to the number of these independent events, the creations of molecules destined to survive until t, during the period $[0,t]$. Therefore, the molecule number follows the Poisson distribution $P_{\rm Poisson}(n;\mu(t))$ with

$$\mu(t) = \int_0^t p_{\rm surv}(t|s)\alpha(s)ds = \int_0^t \exp\left(-\int_s^t \beta(u)du\right)\alpha(s)ds. \tag{22}$$

The master equation for this model is

$$\frac{\partial P(n,t)}{\partial t} = \alpha(t)\left[P(n-1,t) - P(n,t)\right] + \beta(t)\left[(n+1)P(n+1,t) - nP(n,t)\right], \tag{23}$$

and it is straightforward to check by direct substitution that $P_{\rm Poisson}(n;\mu(t))$ with $\mu(t)$ given by Eq.(22) is a solution of this equation.

We noted in the previous section that the distribution of the molecule number in the degradation-only model is Poissonian for $t > 0$ if the initial distribution at $t = 0$ is also Poissonian. In fact, this distribution can be considered as a special case of the Poisson distribution in the model with a time-dependent creation rate $\alpha(t)$, by shifting the origin of time: We start from zero molecules at some time $t_0 < 0$ so that we obtain a Poisson distribution for $P(n, 0)$. We then require that $\alpha(t) = 0$ for $t > 0$ so that the model reduces to the degradation-only model for $t > 0$.

The probability distribution converges to a stationary Poisson distribution if and only if the expected number of surviving molecules at $t \to \infty$, given by the integral

$$\mu(\infty) = \int_0^\infty \alpha(s) \exp\left(-\int_s^\infty \beta(u)du\right) ds, \tag{24}$$

is finite. For example, if the rates α and β are constants, we get

$$\mu(t) = \int_0^t \alpha e^{-(t-s)\beta} ds = \frac{\alpha}{\beta}(1 - e^{-\beta t}), \tag{25}$$

and

$$\mu(\infty) = \frac{\alpha}{\beta}. \tag{26}$$

Now consider a more general situation where the initial number of molecules is n_0. The number of molecules at a later time t is the sum of the number n_1 of surviving molecules among initial n_0 particles, whose distribution follows the binomial distribution in Eq.(15), and the number n_2 of surviving molecules created during the interval $[0, t]$, which follows the Poisson distribution. Therefore, the probability distribution $P(n, t)$ for the molecule number n is expressed as a convolution of the binomial distribution and the shifted Poisson distribution:

$$\begin{aligned} P(n,t|n_0) &= \sum_{n_1+n_2=n} P_{\rm binom}(n_1; \{n_0, p_{\rm surv}(t|0)\}) P_{\rm Poisson}(n_2; \mu(t)) \\ &= \sum_{n_1} \frac{n_0!}{n_1!(n_0-n_1)!} \frac{e^{-\mu(t)}\mu(t)^{n-n_1}}{(n-n_1)!} p_{\rm surv}(t|0)^{n_1}(1 - p_{\rm surv}(t|0))^{n_0-n_1}, \end{aligned} \tag{27}$$

where $p_{\rm surv}(t|0)$ and $\mu(t)$ are given by Eq.(14) and Eq.(22), respectively. Here, $\mu(t)$ is not the expected number of molecules at t anymore. $\mu(t)$ is the expected number of surviving molecules created within the time interval $[0,t]$, and one has to add the expected number of surviving ones among the initial n_0 molecules in order to get the expected total number of molecules, which I denote as $\mathrm{E}[n|n^0]$:

$$\begin{aligned}
\mathrm{E}[n|n^0] &\equiv \sum_n nP(n,t|n_0) = \sum_{n_1,n_2}(n_1+n_2)P_{\rm binom}(n_1;\{n_0,p_{\rm surv}(t|0)\})P_{\rm Poisson}(n_2;\mu(t)) \\
&= \sum_{n_1} n_1 P_{\rm binom}(n_1;\{n_0,p_{\rm surv}(t|0)\}) + \sum_{n_2} P_{\rm Poisson}(n_2;\mu(t)) \\
&= n_0 p_{\rm surv}(t|0) + \mu(t) \\
&= n_0 \exp\left(-\int_0^t \beta(s)ds\right) + \int_0^t \exp\left(-\int_s^t \beta(u)du\right)\alpha(s)ds
\end{aligned} \tag{28}$$

It is straightforward to check that the distribution in (27) is a solution of the master equation (23) by direct substitution (Appendix C). Eq.(27) is a special case of the time-dependent distribution for a more general multispecies monomolecular reaction network [29], whose generalized versions will be discussed later in sections VII and VIII.

By the linearity of the master equation, the solution for an arbitrary initial distribution $P(n,0)=v(n)$ can be constructed by superposing the distributions in Eq.(27),

$$\begin{aligned}
P(n,t) &= \sum_{n_0} P(n,t|n_0)v(n_0) \\
&= \sum_{n_0,n_1} \frac{v(n_0)n_0!e^{-\mu(t)}\mu(t)^{n-n_1}}{n_1!(n_0-n_1)!(n-n_1)!} p_{\rm surv}(t|0)^{n_1}(1-p_{\rm surv}(t|0))^{n_0-n_1}.
\end{aligned} \tag{29}$$

The analytic expression for $P(n,t|n_0)$ in Eq.(27) is plotted with solid lines for time points $\beta t = 0.1$, 0.5, and 2.0 in Figure 1, for $\alpha = 5.0\beta$ and $n_0 = 10$, where the distribution converges to a Poisson distribution with $\mu(\infty) = \frac{\alpha}{\beta} = 5.0$. The results obtained by Gillespie stochastic simulation [35] are also shown with filled circles, which agree almost perfectly with the analytic formula.

VI. MODEL WITH TIME-DELAYED DEGRADATION

Due to the complex mechanism of protein degradation, there can be a time delay in protein degradation. A model with time-delayed degradation attempts to capture such behavior, and once a molecule enters the degradation process, it gets degraded after a fixed

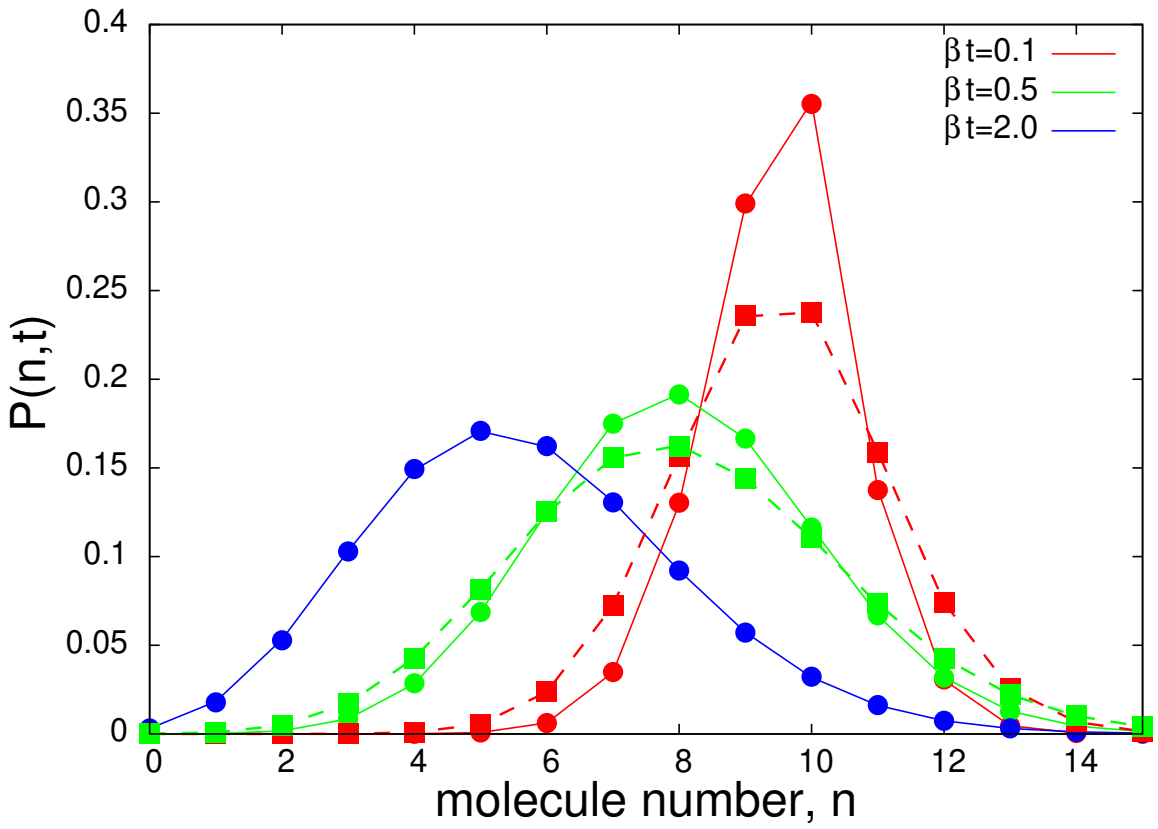

FIG. 1. The probability distribution $P(n,t|n_0)$ of the molecule numbers for the model with molecule creation and degradation, obtained using the analytic formula Eq.(27) for $\beta t = 3.0$, 5.0 and 7.0 (solid lines), with the initial molecule number $n_0 = 10$ and the parameter $\alpha = 5.0\beta$. The dashed lines are the results for the two species model with $n_1^0 = n_2^0 = 10$, $\beta_1 = \beta_2 = \beta$, $\alpha_1 = \alpha_2 = 5.0\beta$, and $c_{12} = c_{21} = \beta$, obtained using the analytic formula Eq.(52). The symbols on top of the lines are the results of the stochastic simulations, where 10^6 independent simulations were averaged. The dependence on the initial molecule numbers is omitted in the vertical label for notational simplicity.

time delay τ [16]. A more general model where the molecule is allowed to get degraded before τ was also constructed [17], which I will examine in this section in detail. The model

is

$$\varnothing \xrightarrow{\alpha} X_A, \quad X_A \xrightarrow{\gamma} \varnothing, \quad X_A \xrightarrow{\beta} X_I, \quad X_I \underset{\tau}{\Rightarrow} \varnothing, \quad X_I \xrightarrow{\zeta} \varnothing. \tag{30}$$

Here, X_I is an inactive molecule that has entered the degradation process. An active molecule is denoted by X_A, which can undergo instantaneous degradation with rate γ. It can enter the delayed degradation process with rate β, at which point it becomes inactive and gets degraded after time τ with certainty, but it can also undergo instantaneous degradation with rate ζ before the delayed degradation process is complete. Although it is straightforward to allow for time-dependent rates, I will keep them time-independent in order to compare the result with that of ref. [17], as well as for notational simplicity.

This model is non-Markovian, and the master equation for the probability distribution $P(n_A, n_I, t)$ takes the form

$$\begin{aligned}\frac{dP(n_A,n_I,t)}{dt} &= (E_A^- - 1)\alpha P(n_A,n_I,t) + (E_A^+ - 1)\gamma n_A P(n_A,n_I,t) \\ &+ (E_A^+ E_I^- - 1)\beta n_A P(n_A,n_I,t) + (E_I^+ - 1)\zeta n_I P(n_A,n_I,t) \\ &+ \sum_{n'_A} P^*(n_A, n_I - 1, \tau | n'_A - 1)\beta n'_A P(n'_A, t-\tau) e^{-\zeta\tau},\end{aligned} \tag{31}$$

where $E^{\pm}_{A,I} f(n_{A,I}) \equiv f(n_{A,I} \pm 1)$, $P(n_A, t) \equiv \sum_{n_I} P(n_A, n_I, t)$ is the marginal probability for n_A, and $P^*(n_A, n_I, t | n'_A - 1)$ is the probability under the initial condition of $n'_A - 1$ active molecules and no inactive molecules, obtained by neglecting the time-delayed degradation [17]:

$$\begin{aligned}\frac{dP^*(n_A,n_I,t|n'_A-1)}{dt} &= (E_A^- - 1)\alpha P^*(n_A,n_I,t|n'_A-1) + (E_A^+ - 1)\gamma n_A P^*(n_A,n_I,t|n'_A-1) \\ &+ (E_A^+ E_I^- - 1)\beta n_A P^*(n_A,n_I,t|n'_A-1) \\ &+ (E_I^+ - 1)\zeta n_I P^*(n_A,n_I,t|n'_A-1).\end{aligned} \tag{32}$$

Eq. (31) along with Eq. (32) admits the Poisson distribution as a time-dependent solution if the particle number is initially zero or the initial distribution is Poissonian [17]. However, it is easy to derive the Poisson distribution without going through the complicated process of solving Eqs. (31) and (32), by considering the underlying Poisson process.

Note that one of three things happen in a given short time interval of $[t', t' + dt']$: either a molecule that will survive at t as an active one is created, the one that will survive as an inactive one is created[2], or none of these happens. These events are independent of

[2] This should not be confused with a birth of an inactive molecule. Every molecule is born active, and only an active molecule can turn into an inactive one. Here I am considering a birth of an active molecule that will eventually survive as an inactive molecule at the later time t

creation events in other regions in time. Also, although the creation of a surviving active molecule and a surviving inactive molecule are exclusive events, they can be treated as independent events because they are rare events: even if we assume they are independent, the probability of more than two molecules created in $[t', t' + dt']$ is negligible. Therefore, exclusive and independent events are indistinguishable if they are rare (Appendix D), and the probability distribution for the numbers n_A, n_I of X_A, X_I are given by the product of Poisson distributions,

$$P(n_A, n_I, t|0) = \frac{e^{-\mu_A(t)-\mu_I(t)}}{n_A! n_I!} \mu_A(t)^{n_A} \mu_I(t)^{n_I}. \tag{33}$$

It only remains to compute $\mu_A(t)$ and $\mu_I(t)$.

Suppose that a molecule is created at time t'. For the molecule to remain active at a later time t, (i) it should not undergo instantaneous degradation ($X_A \xrightarrow{\gamma} \varnothing$), and (ii) should not turn into an inactive molecule ($X_A \xrightarrow{\beta} X_I$), during the intervening time. Therefore, the conditional probability $p_A(t - t')$ that this molecule will remain active at t is given by the exponential distribution in $t - t'$,

$$p_A(t - t') = e^{-a(t-t')}, \tag{34}$$

where $a \equiv \beta + \gamma$, and we get

$$\mu_A(t) = \int_0^t \alpha p_A(t - t') dt' = \frac{\alpha}{a} \left(1 - e^{-at}\right). \tag{35}$$

For a molecule created at time t' to survive at a later time t as an inactive molecule, (i) it must turn into an inactive one within a short time interval $[t_I, t_I + dt_I]$ such that $\max(t', t - \tau) < t_I < t$, (ii) it should not undergo instantaneous degradation ($X_A \xrightarrow{\gamma} \varnothing$) in the time between t' and t_I, and (iii) it should not undergo instantaneous degradation ($X_I \xrightarrow{\zeta} \varnothing$) in the time between t_I and t. For a given value of t_I, such a conditional probability $\rho_I(t, t_I|t')dt_I$ is given as

$$\rho_I(t, t_I|t')dt_I = e^{-a(t_I-t')} e^{-\zeta(t-t_I)} \beta dt_I \tag{36}$$

Integrating over t_I, we get the conditional probability $p_I(t - t')$ that a molecule created at time t' will survive as an inactive one at t:

$$\begin{aligned} p_I(t - t') &= \int_{\max(t', t-\tau)}^{t} \rho_I(t, t_I|t')dt_I = \frac{\beta e^{at'-\zeta t}}{\zeta - a} \left(e^{(\zeta-a)t} - e^{(\zeta-a)\max(t', t-\tau)}\right) \\ &= \begin{cases} \beta(\zeta - a)^{-1} \left(e^{a(t'-t)} - e^{\zeta(t'-t)}\right) & (0 \le t - t' \le \tau), \\ \beta(\zeta - a)^{-1} e^{a(t'-t)} \left(1 - e^{(a-\zeta)\tau}\right) & (\tau \le t - t' < t). \end{cases} \end{aligned} \tag{37}$$

The expected number of surviving inactive molecules, $\mu_I(t)$, is then obtained by multiplying $p_I(t-t')$ by the creation rate α and integrating over t':

$$\mu_I(t) = \int_0^t \alpha p_I(t-t')dt' = \begin{cases} \frac{\alpha\beta}{\zeta-a}\left[\frac{1}{a}(1-e^{-at}) - \frac{1}{\zeta}(1-e^{-\zeta t})\right] & (0 \le t < \tau), \\ \frac{\alpha\beta}{a}\left[\frac{1-e^{-\zeta\tau}}{\zeta} + \frac{1-e^{(a-\zeta)\tau}}{a-\zeta}e^{-at}\right] & (t \ge \tau). \end{cases} \tag{38}$$

Eq.(33) along with Eq.(35) and Eq.(38) completely specify the time-dependent Poisson distribution when we start from zero molecules at $t = 0$, which agrees with the result in ref. [17]. The total number of molecules, $n = n_A + n_I$, also follows a Poisson distribution with the expectation value $\mu(t) = \mu_A(t) + \mu_I(t)$, because the sum of two variables that follow Poisson distributions also follows a Poisson distribution [33]. The Poisson process underlying the Poisson distribution for $n = n_A + n_I$ is the creation of particles destined to survive until time t, regardless of being active or inactive.

Finally, as in the case of the Markovian model in the previous section, we can derive general time-dependent distributions by combining the Poisson distribution with the multinomial distribution. However, note that it is almost impossible to allow arbitrary initial distribution at $t = 0$. If a non-zero initial value of n_I is allowed, there is an ambiguity in the evolution of the system for $t > 0$ because the final degradation of the initial inactive molecules depends on the exact time points at $t < 0$ that they got inactivated. Therefore, I will only consider the case where only active particles are present at $t = 0$, whose number is n^0. Given an active molecule at $t = 0$, the probability that it will survive as an active molecule and the probability that it will survive as an inactive molecule, at a later time t, are given by Eqs. (34) and (37), respectively, with $t' = 0$:

$$\begin{aligned} p_A(t) &= e^{-at}, \\ p_I(t) &= \begin{cases} \beta(\zeta-a)^{-1}\left(e^{-at} - e^{-\zeta t}\right) & (t < \tau), \\ \beta(\zeta-a)^{-1}e^{-at}\left(1 - e^{(a-\zeta)\tau}\right) & (t \ge \tau). \end{cases} \end{aligned} \tag{39}$$

Therefore, the probability that there are n'_A surviving active molecules and n'_I surviving active molecules *among* initial n_0 active molecules, at time t, is given by the multinomial distribution,

$$\begin{aligned} P_{\text{mult}}(n'_A, n'_I; \{n_0, p_A(t), p_I(t)\}) \equiv \frac{n_0!}{n'_A! n'_I! (n_0 - n'_A - n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\ \times (1 - p_A(t) - p_I(t))^{n_0 - n'_A - n'_I}. \end{aligned} \tag{40}$$

The numbers of active and inactive molecules at t are decomposed as $n_A = n'_A + n''_A$, $n_I = n'_I + n''_I$, where $n''_A$and n''_I are the numbers of the surviving active and inactive molecules created in the time interval $[0, t]$. Therefore, the probability distribution for n_A and n_I at t is given by the convolution of the multinomial and the Poisson distribution:

$$\begin{aligned} P(n_A, n_I, t|n_0) = & \sum_{n'_A+n''_A=n_A} \sum_{n'_I+n''_I=n_I} P_{\text{mult}}(n'_A, n'_I; \{n_0, p_A(t), p_I(t)\}) \\ & \times P_{\text{Poisson}}(n''_A; \mu_A(t)) P_{\text{Poisson}}(n''_I; \mu_I(t)) \\ = & \sum_{n'_A} \sum_{n'_I} \frac{n_0!}{n'_A! n'_I! (n_0 - n'_A - n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\ & \times (1 - p_A(t) - p_I(t))^{n_0 - n'_A - n'_I} \\ & \times \frac{e^{-\mu_A(t) - \mu_I(t)} \mu_A(t)^{n_A - n'_A} \mu_I(t)^{n_I - n'_I}}{(n_A - n'_A)!(n_I - n'_I)!}. \end{aligned} \tag{41}$$

We can check that the distribution in Eq.(41) is indeed a solution of the master equation Eq.(31) by direct substitution (Appendix E). Again, by the linearity of the master equation, the time-dependent probability distribution for an arbitrary initial distribution of *active* molecules,

$$P(n^0_A, n^0_I, 0) = v(n^0_A) \delta_{n^0_I, 0}, \tag{42}$$

is obtained by the superposition of Eq.(41) weighted by $v(n_0)$,

$$\begin{aligned} P(n_A, n_I, t) = \sum_{n_0} P(n_A, n_I, t|n_0) v(n_0) = & \sum_{n_0} \sum_{n'_A} \sum_{n'_I} v(n_0) \frac{n_0!}{n'_A! n'_I! (n_0 - n'_A - n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\ & \times (1 - p_A(t) - p_I(t))^{n_0 - n'_A - n'_I} \\ & \times \frac{e^{-\mu_A(t) - \mu_I(t)} \mu_A(t)^{n_A - n'_A} \mu_I(t)^{n_I - n'_I}}{(n_A - n'_A)!(n_I - n'_I)!}. \end{aligned} \tag{43}$$

VII. MULTISPECIES MONOMOLECULAR REACTION THAT INCLUDES SPECIES WITH FINITE UPPER BOUNDS ON LIFETIMES

Is there a common feature of the models considered in the previous sections, which leads to the Poissonian distribution? In fact, there is, which is that these are examples of monomolecular reaction networks [29, 36, 37]. A monomolecular reaction network is a set of reactions where neither the number of reactants nor that of the products exceeds one for each reaction. It has been shown that for a multispecies monomolecular reaction network of the form

$$X_i \xrightarrow{c_{ij}(t)} X_j \quad (i \neq j), \quad \varnothing \xrightarrow{\alpha_i(t)} X_i, \quad X_i \xrightarrow{\beta_i(t)} \varnothing, \tag{44}$$

for species $X_1, \cdots, X_\ell$, the most general time-dependent probability distribution of the molecule numbers is obtained by superposing the convolutions of the multinomial distributions with the Poisson distributions [29], which is the multi-species generalization of the expression Eq.(27). In this section, I derive a generalized version of this solution using the viewpoint of "counting of independent events", for a monomolecular reaction network that includes species with finite upper bounds on their lifetimes, so that the model with time-delayed degradation can be included.

It is clear from the discussions in the previous sections why monomolecularity is so important for the appearance of the Poisson distribution. To derive the Poisson distribution, the number of molecules at a given time should be interpretable as the number of independent creation events that occurred during some time interval in the past. For a monomolecular reaction network, the creations of each molecule are independent of each other. Subsequent degradations or conversions of these molecules are independent of each other as well. A molecule can disintegrate, but it does not get split into more than one molecule, and consequently, we can talk about a creation event of a molecule that will eventually survive at a later time as a particular species X_i. This would be difficult if the number of reactants or products of some reaction is greater than two, where the fates of different molecules get entangled with each other.

The model in Eq.(44) is generalized to include species with finite upper bounds on their lifetimes, so that it encompasses the model with time-delayed degradation discussed in the previous section, by including additional reactions $X_i \underset{\tau_i}{\Longrightarrow} \varnothing$ with $\tau_i \in [0, \infty]$, so that we have

$$X_i \xrightarrow{c_{ij}(t)} X_j \quad (i \neq j), \quad \varnothing \xrightarrow{\alpha_i(t)} X_i, \quad X_i \xrightarrow{\beta_i(t)} \varnothing, \quad X_i \underset{\tau_i}{\Longrightarrow} \varnothing. \tag{45}$$

That is, there is an upper limit τ_i that species X_i can persist without transforming to other species, after which it must disintegrate. For a species X_i without a finite bound on its lifetime, we can simply set $\tau_i = \infty$. Then the delayed degradation model in the previous section corresponds to the case with $\ell = 2$, $\alpha_2 = c_{21} = 0$, and $\tau_1 = \infty$, whereas the model in Eq.(44) corresponds to the case of $\tau_i = \infty$ for all i. Just as in the previous sections, we note that if the number of molecules has been zero for $t \leq 0$, the number of X_i molecules at a later time t is again equal to the number of creation events of molecules during the time interval $[0, t]$, which are destined to survive as X_i at the end. Since these are independent events, we deduce that the joint probability distribution for the molecule numbers at time

t is the product of the Poisson distributions

$$P(n_1, \cdots, n_\ell; t|\mathbf{0}) = \prod_{i=1}^{\ell} \frac{e^{-\mu_i}}{n_i!} \mu_i^{n_i} \tag{46}$$

where μ_i is the expected number of molecule creations during $[0, t]$ that are destined to survive as X_i at t. The probability that such a molecule is created as X_j during the time interval $[t', t' + dt']$ is the product of the probability $\alpha_j(t')dt'$ that a molecule of X_j is created in the interval and the probability $p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau})$ that it will survive until time t as X_i. The expected number μ_i of these creation events are then obtained after the summation over j and the integration over t',

$$\mu_i = \sum_j \int_0^t p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau}) \alpha_j(t') dt' \tag{47}$$

Note that in general, the probability that a given molecule at time t' will survive at a later time depends on the history of that molecule. Therefore $p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau})$ is specifically defined as the survival probability conditioned on the *creation* of X_j molecule at time t'. Also, the dependence of $p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau})$ on $\boldsymbol{\tau} \equiv (\tau_1, \cdots, \tau_\ell)$ is explicitly shown to emphasize that the probability is conditioned on $\boldsymbol{\tau}$ as well as t'.

$p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau})$ can be obtained by considering its change with time. As elaborated in Appendix F, the equation describing the time evolution of $p_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau})$ is given by

$$\begin{aligned}
\dot{p}_{ij}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau}) = & \sum_k B_{ik}(t) p_{kj}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau}) \\
& - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_{k \neq i} B_{ik}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t', \boldsymbol{\tau}) \\
& - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \delta_{ij}\delta(t-\tau_i-t')
\end{aligned} \tag{48}$$

where

$$\begin{aligned}
B_{ji}(t) &= c_{ij}(t) \quad (i \neq j) \\
B_{ii}(t) &= -\sum_{j=0}^{\ell} c_{ij}(t) \equiv -\beta_i(t) - \sum_{j=1}^{\ell} c_{ij}(t),
\end{aligned} \tag{49}$$

and $p_{kj}^{(\mathrm{surv})}(t''|t', \boldsymbol{\tau}) \equiv 0$ for $t'' < t'$. The term on the right-hand side of the first line of Eq.(48) is the contribution from the interspecies conversion at time t. The second and the third line come from the disintegration of X_i molecules that have been surviving from

$t-\tau_i$ without transforming to another species, where the second line is the contribution from those converted from another species at $t-\tau_i$, and the third line is the one from X_i molecules that was created at t' when $t'=t-\tau_i$. Eqs.(48) along with the initial condition $p_{ij}^{(\mathrm{surv})}(t'|t',\boldsymbol{\tau})=\delta_{ij}$ fully specifies $p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})$ for $t\geq t'$. The equation for μ_i is then obtained by taking the time derivative of Eq.(47) and using Eq. (48),

$$\begin{aligned}\dot{\mu}_i(t|\boldsymbol{\tau}) &= \alpha_i(t)+\sum_k B_{ik}(t)\mu_k(t|t',\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right)\sum_{k\neq i} B_{ik}(t-\tau_i)\mu_k(t-\tau_i|t',\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right)\alpha_i(t-\tau_i) \end{aligned} \tag{50}$$

where $\alpha_i(t'')=\mu_i(t'')\equiv 0$ for $t''<0$. The detailed steps for the derivation of Eq.(50) are also presented in Appendix F. Eq.(50) along with the initial condition $\mu_i(0|\boldsymbol{\tau})=0$ fully specifies $\mu_i(t|\boldsymbol{\tau})$ for $t\geq 0$, which in turn fully defines the time-dependent probability distribution Eq. (46) of molecules numbers for the system in Eq.(45), when the numbers of the molecules are zero for $t\leq 0$.

Without loss of generality, let us assume that $\tau_i=\infty$ for $1\leq i\leq m$, with $0\leq m\leq \ell$. Then we can consider a more general initial condition with nonzero initial numbers of X_i molecules for $i=1,\cdots m$, as was done in the previous section for the model with a time-delayed degradation. First, for a given species X_i $(1\leq i\leq m)$ with the initial number of molecules n_i^0, let us denote the number of molecules that survive until time t as species X_j $(1\leq j\leq \ell)$ as n_{ij}. Then, the number of those that decomposed among the initial n_i^0 molecules is $n_i^0-\sum_{j=1}^{\ell} n_{ij}$, and the probability for $\mathbf{n}_i\equiv(n_{i1},n_{i2},\cdots n_{i\ell})$ is given by the multinomial distribution,

$$\begin{aligned} P_{\mathrm{multi}}(\mathbf{n}_i;n_i^0,\mathbf{p}_i^{(\mathrm{surv})}(t|0,\boldsymbol{\tau})) &\equiv \frac{n_i^0!}{\prod_{j=1}^{\ell} n_{ij}!(n_i^0-\sum_{p=1}^{\ell} n_{ip})!}\prod_{k=1}^{\ell} p_{ik}^{(\mathrm{surv})}(t|0,\boldsymbol{\tau})^{n_{ik}} \\ &\quad\times\left(1-\sum_{q=1}^{\ell} p_{iq}^{(\mathrm{surv})}(t|0,\boldsymbol{\tau})\right)^{n_i^0-\sum_{r=1}^{\ell} n_{ir}}, \end{aligned} \tag{51}$$

where $\mathbf{p}_i^{(\mathrm{surv})}(t|0,\boldsymbol{\tau})\equiv(p_{i1}^{(\mathrm{surv})}(t|0,\boldsymbol{\tau}),p_{i2}^{(\mathrm{surv})}(t|0,\boldsymbol{\tau}),\cdots,p_{i\ell}^{(\mathrm{surv})}(t|0,\boldsymbol{\tau}))$. Denoting the number of X_j $(1\leq j\leq \ell)$ molecules at time t as n_j, we see that $n_j=\sum_{i=0}^{m} n_{ij}$ where n_{0j} is the number of newly created X_j molecules during the time interval $[0,t]$, whose probability distribution was already written down as Eq.(46). Therefore, the time distribution of the

molecule numbers is expressed as a convolution of m multinomial distributions and ℓ Poisson distributions,

$$P(n_1,\cdots,n_\ell;t|n_1^0,\cdots,n_m^0,\boldsymbol{\tau}) = \sum_{\sum_{j=0}^m \mathbf{n}_j=\mathbf{n}} \left(\prod_{i=1}^m P_{\rm multi}(\mathbf{n}_i;n_i^0,\mathbf{p}_i^{\rm (surv)}(t|0,\boldsymbol{\tau}))\right) \times \left(\prod_{k=1}^\ell P_{\rm Poisson}(n_{0k};\mu_k(t))\right). \tag{52}$$

The distribution in Eq.(52) encompasses Eq.(46) as a special case when $n_i^0 = 0$ for all i. As in the case of the model with delayed degradation considered in the previous section, I do not consider more general initial conditions with nonzero numbers of species with finite upped bounds of their lifetimes: For those species, the degradation probability of a molecule depends on the history and therefore the evolution of the probability distribution is not determined solely by the distribution at $t = 0$. As in the previous sections, for an arbitrary distribution $v(n_1^0,\cdots v_m^0)$ of molecules of species $X_1,\cdots,X_m$, the probability distribution of the molecules numbers at a later time is given by

$$P(n_1,\cdots,n_\ell;t|\boldsymbol{\tau}) = \sum_{n_1^0,\cdots n_m^0} P(n_1,\cdots,n_\ell;t|n_1^0,\cdots,n_m^0,\boldsymbol{\tau})v(n_1^0,\cdots,n_m^0), \tag{53}$$

which is a generalization of Eq.(43). When $m = \ell$, the expression for $P(n_1,\cdots,n_\ell;t|n_1^0,\cdots,n_m^0,\boldsymbol{\tau})$ in Eq.(52) reduces to that for the model in Eq.(44) without finite upper bounds on molecule lifetimes, as detailed in Appendix G. In this case, I will drop the $\boldsymbol{\tau}$ dependence altogether and simply write $P(n_1,\cdots,n_\ell;t|n_1^0,\cdots,n_\ell^0,\infty,\cdots\infty)$ as $P(n_1,\cdots,n_\ell;t|n_1^0,\cdots,n_\ell^0)$. The probability distribution $P(n_1,n_2;t|10,10)$ is shown in Fig. 1 as dashed lines for the model with $\ell = m = 2$ and $c_{12} = c_{21} = \beta \equiv \beta_1 = \beta_2$, for the timepoints at $\beta t = 0.1$ and $\beta t = 0.5$. Since the distribution converges to a Poisson distribution with $\mu(\infty|10,10) = 5.0$, the graph for the $\beta t = 2.0$ almost overlaps with the solid line and therefore was not plotted. Although the initial number of molecules for each of the species X_1 and X_2, as well as the stationary distribution, are chosen to be the same as those for the one-species model plotted by the solid lines, we find that the transient behavior of the distribution is different because of the interspecies conversion. The results of the Gillespie simulation [35] are also shown with filled squares, which show nearly perfect agreement with the analytic formula.

VIII. MULTISPECIES MONOMOLECULAR REACTION THAT INCLUDES SPECIES WITH NON-EXPONENTIAL DISTRIBUTIONS OF LIFETIMES

Now I generalize the model in Eq.(45) one step further, and consider a case where the lifetime τ_i of a X_i molecule created at time t is a stochastic variable that follows a general time-dependent distribution $f_i(\tau_i|t)$. This model can be written as

$$X_i \xrightarrow{c_{ij}(t)} X_j \quad (i \neq j), \quad \varnothing \xrightarrow{\alpha_i(t)} X_i, \qquad X_i \xRightarrow[f(\tau_i|t)]{} \varnothing. \tag{54}$$

The model in Eq.(54) includes the model with a time-delayed degradation with stochastic delay time [16] as a special case. The independence of the creation events remains unchanged, so the probability distribution of the molecule numbers is still of the form Eq.(46) when the molecule numbers vanish for $t \leq 0$, only with the change of the expressions for μ_i. First, the survival probabilities are now obtained as

$$\mathbf{p}^{(\mathrm{surv})}(t|t') = \int \cdots \int \mathbf{p}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau}) \prod_{k=1}^{\ell} f_k(\tau_k|t') d\tau_k, \tag{55}$$

and therefore

$$\boldsymbol{\mu}(t) = \int_0^t \left[\prod_{k=1}^{\ell} \int f_k(\tau_k|t') d\tau_k \mathbf{p}^{(\mathrm{surv})}(t|t', \boldsymbol{\tau}) \right] \boldsymbol{\alpha}(t') dt'. \tag{56}$$

When the distributions $f_k(\tau_k)$ are time-independent, Eq.(56) is further simplified as

$$\boldsymbol{\mu}(t) = \prod_{k=1}^{\ell} \int f_k(\tau_k) d\tau_k \boldsymbol{\mu}(t|\boldsymbol{\tau}). \tag{57}$$

Note that we do not have to treat the process $X_i \xrightarrow{\beta_i(t)} \varnothing$ separately in Eq.(54), since it can be incorporated in $X_i \xRightarrow[f(\tau_i|t)]{} \varnothing$ by using $f_i(\tau_i|t) = \beta_i(t+\tau_i) \exp\left(- \int_t^{t+\tau_i} \beta_i(u) du \right)$. The crucial generalization in the model in Eq.(54) is that non-exponential function is allowed as $f(\tau|t)$, of which the exponential distribution along with finite upper bound considered in Eq.(45) is a special case.

When $f_i(\tau_i|t) = \beta(t+\tau_i) \exp\left(- \int_t^{t+\tau_i} \beta(u) du \right)$ for $i = 1, \cdots m$ $(m \leq \ell)$, then we can consider a more general initial condition with nonzero molecule numbers $n_1^0, \cdots, n_m^0$ for $X_1, \cdots X_m$. Using the same logic as in the previous section, we find that the corresponding

probability distribution $P(n_1, \cdots, n_\ell; t|n_1^0, \cdots, n_m^0)$ at a later time t is again given by

$$P(n_1, \cdots, n_\ell; t|n_1^0, \cdots, n_m^0) = \sum_{\sum_{j=0}^m \mathbf{n}_j = \mathbf{n}} \left(\prod_{i=1}^m P_{\rm multi}(\mathbf{n}_i; n_i^0, \mathbf{p}_i^{(\rm surv)}(t|0)) \right) \times \left(\prod_{k=1}^\ell P_{\rm Poisson}(n_{0k}; \mu_k(t)) \right). \tag{58}$$

with $\mathbf{p}_i^{(\rm surv)}(t|0)$ and $\boldsymbol{\mu}(t)$ given by Eqs.(55) and (56). This is the probability distribution expressed as a convolution of the multinomial distribution with the Poisson distribution in the most general settings discussed in the current work, that encompasses all the results in the previous sections as well as those in the literature [16, 17, 28, 29, 36] as special cases.

The average and the variance of the molecule number are found from the well-known results for multinomial and Poisson variables:

$$\begin{aligned} {\rm E}[n_i|n_1^0, \cdots, n_m^0](t) &\equiv \sum_{n_1, \cdots n_k} n_i P(n_1, \cdots, n_\ell; t|n_1^0, \cdots, n_m^0) = \mu_i(t) + \sum_{k=1}^m n_k^0 p_{ik}^{(\rm surv)}(t|0) \\ {\rm Var}[n_i|n_1^0, \cdots, n_m^0](t) &\equiv {\rm E}\left[\Big(n_i - {\rm E}[n_i|n_1^0, \cdots, n_m^0](t) \Big)^2 \middle| n_1^0, \cdots, n_m^0 \right](t) \\ &= \mu_i(t) + \sum_{k=1}^m n_k^0 p_{ik}^{(\rm surv)}(t|0)(1 - p_{ik}^{(\rm surv)}(t|0)). \end{aligned} \tag{59}$$

${\rm E}[n_i|n_1^0, \cdots, n_m^0](t)$ and ${\rm Var}[n_i|n_1^0, \cdots, n_m^0](t)$ for the model with $\ell = 3$, $m = 2$, $f_1(\tau_1) = f_2(\tau_2) = \beta e^{-\beta\tau}$, $f_3(\tau_3) = 0.5\delta(\tau_3 - \beta^{-1}) + 0.5\delta(\tau_3 - 2\beta^{-1})$, $\alpha_1 = \alpha_2 = c_{12} = c_{21} = c_{23} = \beta$, $\alpha_3 = c_{3i} = 0$, are plotted as the function of βt in Fig. 2, for the initial condition of $n_1^0 = n_2^0 = 2$. This is the model where the species X_1 and X_2 are created with the same rate and freely convert to each other, but X_1 goes through usual exponential degradation whereas the X_2 goes through delayed degradation by first transforming to X_3, which will decompose after either $1.0\beta^{-1}$ or $2.0\beta^{-1}$ with equal probabilities. The parameters and the initial conditions are chosen so that X_1 and X_2 exhibit the same behaviors. We see discontinuity in the derivative of ${\rm E}[n_3|2,2](t)$ at the two delay times $\beta t = 1.0$ and $\beta t = 2.0$, because of the nonzero initial numbers of X_2. Half of the X_3 molecules that began to be converted from X_2 molecules at $t = 0$ begin to disintegrate at $t = 1.0\beta$ due to the exhaustion of their lives, and the remaining half begin to die at $t = 2.0\beta$. Such discontinuities in the derivative would be absent if we had $n_0^2 = 0$. Initially, ${\rm E}[n_i|2,2](t)$ and ${\rm Var}[n_i|2,2](t)$ differ, but they converge at later times as the effect of the initial molecules dies out and the distributions approach Poisson distributions. The simulation result using Cai's extended version [38] of

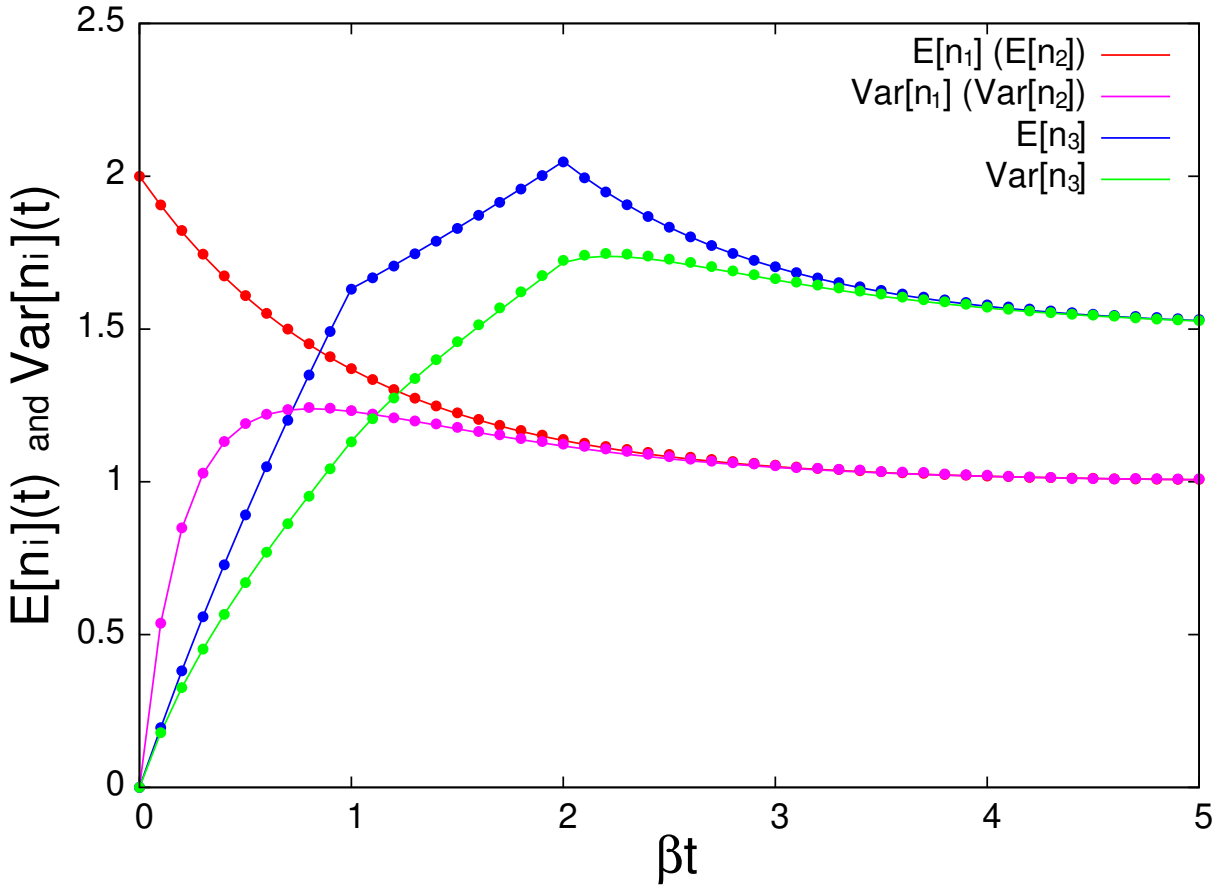

FIG. 2. $\mathrm{E}[n_i|2,2](t)$ and $\mathrm{Var}[n_i|2,2](t)$ as the function of βt, for the model with $f_1(\tau_1) = f_2(\tau_2) = \beta e^{-\beta\tau}$, $f_3(\tau_3) = 0.5\delta(\tau_3 - \beta^{-1}) + 0.5\delta(\tau_3 - 2\beta^{-1})$, $\alpha_1 = \alpha_2 = c_{12} = c_{21} = c_{23} = \beta$, $\alpha_1 = c_{3i} = 0$. The filled circles are the results of stochastic simulations, where the results of 10^6 independent simulations were averaged. The dependence on the initial molecule numbers is omitted in the figure legend and the vertical label for notational simplicity.

the Gillespie algorithm, where delayed reactions are allowed, is plotted with filled circles, again exhibiting nearly perfect agreement with the analytic formula.

IX. MODEL WITH STOCHASTIC RATES

Various external noise sources can be modeled by treating rates themselves as stochastic variables [23, 25, 26, 39]. These models can be used to emulate the regulation of the expres-

sion by the transcription factor binding [25, 39], or the extrinsic noise due to heterogeneous cellular environments [25]. When the creation rate is a stochastic variable, the creations of the molecules form a Poisson process only if the creation rate has no temporal correlations, which is not valid for most of the non-trivial models. However, for a *given* realization of the history of such a stochastic creation rate, we get a creation rate $\alpha(t)$ with a fixed time-dependence, and we obtain a Poisson process as the probability distribution of the molecule numbers, for a monomolecular reaction network. Therefore, a monomolecular reaction model with stochastic rates can still be obtained by from the probability distribution considered in the previous section, by taking a weighted average with respect to the values of μ. As a concrete example, I consider a one-species model with transcriptional pulsing, also called the telegraph process [25, 39, 40].

$$D \underset{k_b}{\overset{k_f}{\rightleftharpoons}} D^*, \quad D^* \xrightarrow{\alpha} D^* + X, \quad X \xrightarrow{\beta} \varnothing, \tag{60}$$

In this model, the gene is transcribed only when it is in the active state, denoted by D^*. Since the transition between the active and the inactive states of the gene follows stochastic dynamics, the model can be considered as a model where the production rate itself is a stochastic variable. For a *given* path of $D \underset{k_b}{\overset{k_f}{\rightleftharpoons}} D^*$ transition, the model reduces to the one with a predetermined time-dependent production rate $\alpha(t)$. Therefore, if the initial number of molecules is zero, the distribution $P(n,t)$ is given by the Poisson distribution with $\mu(t)$ given by Eq.(22). The stochastic dynamics of the gene is incorporated by taking the average with respect to the path probabilities of the gene, and since each path leads to a Poisson distribution with its own parameter, the final expression is a weighted average of the Poisson distributions with distinct parameter values

$$P(n,t) = \int_0^\infty \rho(\mu,t) \frac{e^{-\mu}}{n!} \mu^n d\mu, \tag{61}$$

where $\rho(\mu,t)$ is the probability distribution for the Poisson parameter μ at time t, obtained from the stochastic dynamics of gene [40]. According to the formalism presented in the previous sections, the distribution for arbitrary initial distribution is obtained by convoluting the shifted Poisson distribution with the binomial distribution and then taking the weighted average with respect to the initial distribution, for a *given* value of μ. Therefore, after taking the weighted average over μ values, the general time-dependent distribution is

$$P(n,t) = \sum_{n_0,n_1,j} \frac{v_j(n_0) n_0! p_{\rm surv}(t|0)^{n_1} (1-p_{\rm surv}(t|0))^{n_0-n_1}}{n_1!(n_0-n_1)!(n-n_1)!} \int_0^\infty \rho_j(\mu,t) e^{-\mu} \mu^{n-n_1} d\mu, \tag{62}$$

where $v_j(n_0)$ is the probability that the initial gene state is j $(= I, A)$ and the initial number of molecules is n_0, and $\rho_j(\mu, t)$ is the probability density for μ under the condition that the initial gene state is j $(= I, A)$. For the model in Eq. (60), the computation of $\rho(\mu, t)$ simplifies considerably for $\beta = 0$ since μ then depends only on the *total* duration τ of the gene in the active state so that $\mu = \alpha\tau$. In this case $\rho_j(\mu, t) = \alpha^{-1}\tilde{\rho}_j(\tau, t)$ where $\tilde{\rho}_j(\tau, t)$ is the conditional probability density for τ for the given value of t and j.

Let us first consider the contribution to $\tilde{\rho}_I(\tau, t)$ from the path where the transitions occur $2m + 1$ times with an integer $m(= 0, 1, 2, \cdots)$ so that the final state is active. If we denote the dwell times in the inactive states as $t_1^I, t_2^I \cdots t_{m+1}^I$ and $t_1^A, t_2^A \cdots t_{m+1}^A$, respectively, then the contribution from a *given* sequence of t_i^Is and t_i^As is:

$$e^{-k_f t_1^I} k_f e^{-k_b t_1^A} k_b \times \cdots \times e^{-k_f t_{m+1}^I} k_f e^{-k_b t_{m+1}^A} = k_f^{m+1} k_b^m e^{-k_f(t-\tau)} e^{k_b \tau} \tag{63}$$

where we used the fact that $\sum_{i=1}^{m+1} t_i^I = t - \tau$ and $\sum_{i=1}^{m+1} t_i^A = \tau$ to get the second expression. Similarly, if there are $2m + 2$ $(m = 0, 1, 2, \cdots)$ transitions so that the final state is inactive, the corresponding contribution for a given sequence of t_i^Is and t_i^As is

$$e^{-k_f t_1^I} k_f e^{-k_b t_1^A} k_b \times \cdots \times e^{-k_f t_{m+1}^I} k_f e^{-k_b t_{m+1}^A} k_b e^{-k_f t_{m+2}^I} = k_f^{m+1} k_b^{m+1} e^{-k_f(t-\tau)} e^{k_b \tau}. \tag{64}$$

Finally, the probability that the gene does not make a transition during the time interval $[0, t]$ is $e^{-k_f t}$. Also, we have the conditional probability $\text{Prob}(0 \le \tau \le \Delta t | \text{no transition}) = 1$ for any Δt. Therefore, the contribution to $\tilde{\rho}_I(\tau, t)$ from the path with no transition is

$$e^{-k_f t} \delta_+(\tau), \tag{65}$$

where $\delta_+(\tau)$ is the analogue of the dirac delta distribution defined on the set of non-negative numbers, defined by

$$\int_0^\infty \delta_+(\tau) f(\tau) d\tau = f(0). \tag{66}$$

for any function $f(\tau)$. $\tilde{\rho}_I(\tau, t)$ is then obtained by integrating the expressions in Eqs.(63), (64), and (66) over possible values of t_i^Is and t_i^As, and then summing over m, where the

integration is decoupled and becomes multiple integrals of unity,

$$\begin{aligned}
\rho_I(\tau,t) &= e^{-k_f(t-\tau)}e^{-k_b\tau}\sum_{m\geq 0}k_f^{m+1}k_b^m\int_0^\tau\int_0^{\tau-t_1^A}\cdots\int_0^{\tau-\sum_{i=1}^{m-1}t_i^A}dt_m^A\cdots dt_2^A dt_1^A \\
&\times\int_0^{t-\tau}\int_0^{t-\tau-t_1^I}\cdots\int_0^{t-\tau-\sum_{i=1}^{m-1}t_i^I}dt_m^I\cdots dt_2^I dt_1^I \\
&+ e^{-k_f(t-\tau)}e^{-k_b\tau}\sum_{m\geq 0}k_f^{m+1}k_b^{m+1}\int_0^\tau\int_0^{\tau-t_1^A}\cdots\int_0^{\tau-\sum_{i=1}^{m-1}t_i^A}dt_m^A\cdots dt_2^A dt_1^A \\
&\times\int_0^{t-\tau}\int_0^{t-\tau-t_1^I}\cdots\int_0^{t-\tau-\sum_{i=1}^{m}t_i^I}dt_{m+1}^I\cdots dt_2^I dt_1^I + e^{-k_f t}\delta_+(\tau) \\
&= e^{-k_f t}e^{(k_f-k_b)\tau}\left(\sum_m k_f^{m+1}k_b^m\frac{\tau^m(t-\tau)^m}{m!^2}+\sum_m k_f^{m+1}k_b^{m+1}\frac{\tau^m(t-\tau)^{m+1}}{m!(m+1)!}+\delta_+(\tau)\right)
\end{aligned} \tag{67}$$

$\rho_A(\tau,t)$ is obtained by switching $k_f \leftrightarrow k_b$ and $\tau \leftrightarrow t-\tau$ in $\rho_I(\tau,t)$,

$$\rho_A(\tau,t) = e^{-k_f t}e^{(k_f-k_b)\tau}\left(\sum_m k_f^m k_b^{m+1}\frac{\tau^m(t-\tau)^m}{m!^2}+\sum_m k_f^{m+1}k_b^{m+1}\frac{\tau^{m+1}(t-\tau)^m}{m!(m+1)!}+\delta_+(t-\tau)\right). \tag{68}$$

Therefore, the probability distribution for the molecule number is

$$
\begin{aligned}
P(n,t) &= \sum_{n_0}\sum_{j=I,A}\frac{v_j(n_0)}{(n-n_0)!}\int_0^\infty \rho_I(\mu,t)e^{-\mu}\mu^{n-n_0}d\mu \\
&= \sum_{n_0}\frac{1}{(n-n_0)!}\int_0^t e^{-\alpha\tau}(\alpha\tau)^{n-n_0} \\
&\times \Bigg[v_I(n_0)e^{-k_f t}e^{(k_f-k_b)\tau}\Big(\sum_m k_f^{m+1}k_b^m\frac{\tau^m(t-\tau)^m}{m!^2}+\sum_m k_f^{m+1}k_b^{m+1}\frac{\tau^m(t-\tau)^{m+1}}{(m+1)!m!}+\delta_+(\tau)\Big) \\
&+ v_A(n_0)e^{-k_f t}e^{(k_f-k_b)\tau}\Big(\sum_m k_f^m k_b^{m+1}\frac{\tau^m(t-\tau)^m}{m!^2}+\sum_m k_f^{m+1}k_b^{m+1}\frac{\tau^{m+1}(t-\tau)^m}{(m+1)!m!}+\delta_+(t-\tau)\Big)\Bigg]d\tau \\
&= \sum_{n_0}\frac{(\alpha t)^{n-n_0}}{(n-n_0)!}\Bigg[v_I(n_0)e^{-k_f t} \\
&\times \Big(\sum_m \frac{(k_f t)^{m+1}(k_b t)^m(n-n_0+m)!}{(n-n_0+2m+1)!m!}\Phi(n-n_0+m+1,n-n_0+2m+2;(k_f-k_b-\alpha)t) \\
&+ \sum_m \frac{(k_f t)^{m+1}(k_b t)^{m+1}(n-n_0+m)!}{(n-n_0+2m+2)!m!}\Phi(n-n_0+m+1,n-n_0+2m+3;(k_f-k_b-\alpha)t)\Big) \\
&+ v_A(n_0)e^{-k_f t}\Big(\sum_m \frac{(k_f)^m(k_b t)^{m+1}(n-n_0+m)!}{(n-n_0+2m+1)!m!} \\
&\times \Phi(n-n_0+m+1,n-n_0+2m+2;(k_f-k_b-\alpha)t) \\
&+ \sum_{m\geq 0}\frac{(k_f t)^{m+1}(k_b t)^{m+1}(n-n_0+m+1)!}{(n-n_0+2m+2)!(m+1)!}\Phi(n-n_0+m+2,n-n_0+2m+3;(k_f-k_b-\alpha)t)\Big) \\
&+ v_A(n_0)e^{-(k_b+\alpha)t}\Bigg] + v_I(n)e^{-k_f t} \\
&= \sum_{n_0,n,p,q,r}\frac{v_I(n_0)e^{-k_f t}(\alpha t)^{n-n_0+r}(k_f t)^{p+1}(k_b t)^q(-1)^{q+r}(p+r+n-n_0)!(q+r+n-n_0-1)!}{(n-n_0)!(p+q+r+n-n_0+1)!p!q!r!(r+n-n_0-1)!} \\
&+ \sum_{n_0,n,p,q,r}\frac{v_A(n_0)e^{-k_f t}(\alpha t)^{n-n_0+r}(k_f t)^p(k_b t)^q(-1)^{q+r}(p+r+n-n_0)!(q+r+n-n_0-1)!}{(n-n_0)!(p+q+r+n-n_0)!p!q!r!(r+n-n_0-1)!} \\
&+ v_I(n)e^{-k_f t}
\end{aligned}
\tag{69}
$$

where $\Phi(\alpha,\gamma;z)$ is the degenerate hypergeometric function, also called Kummer's confluent hypergeometric function, defined by

$$
\Phi(\alpha,\gamma;z)\equiv 1+\frac{\alpha}{\gamma}\frac{z}{1!}+\frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!}+\frac{\alpha(\alpha+1)(\alpha+2)}{\gamma(\gamma+1)(\gamma+2)}\frac{z^3}{3!}+\cdots, \tag{70}
$$

and I used the fact that [41]

$$
\int_0^u e^{\gamma x}x^m(u-x)^n=\frac{m!n!}{(m+n+1)!}u^{m+n+1}\Phi(m+1,m+n+2;\gamma u), \tag{71}
$$

to perform the τ integral. The last expression is obtained by using the series representation of the confluent hypergeometric function (Appendix H). The generating function for this model has been obtained in ref. [39] for the special case of $v(n) = \delta_{n,0}$. Although it is difficult to derive the series representation in Eq.(69) starting from the generating function, one can numerically check that it agrees with the generating function obtained from Eq.(69).

Note that for $k_f = k_b = 0$ and $v_I(n) = 0$, the model reduces to the model only with the molecule creation, so the system does not reach a stationary state. However, if $k_b > 0$ and $k_f = 0$, the transcription gets turned off at some point and therefore the average molecule number will not increase indefinitely, and the system will eventually reach a stationary state. One can indirectly see this by computing the average number of molecules at a given time t, assuming that the initial number of molecules was zero. If the transcription was turned off at $t_{\rm off} < t$, then the conditional expectation value is $\langle n(t)\rangle_{t_{\rm off}} = \alpha t_{\rm off}$. On the other hand, if the transcription is turned off at $t_{\rm off} > t$, then we have $\langle n(t)\rangle_{t_{\rm off}} = \alpha t$. Since the probability density $\rho_{\rm off}(t_{\rm off})$ for $t_{\rm off}$ is the exponential distribution $k_b e^{-k_b t_{\rm off}}$, we have

$$\begin{aligned}\mu(t) &= \int_0^\infty \langle n(t)\rangle_{t_{\rm off}} \rho_{\rm off}(t_{\rm off}) = \int_0^t \alpha k_b t_{\rm off} e^{-k_b t_{\rm off}} dt_{\rm off} + \int_t^\infty \alpha k_b t e^{-k_b t_{\rm off}} dt_{\rm off} \\ &= \frac{\alpha}{k_b}(-k_b t e^{-k_b t} + 1 - e^{-k_b t}) + \alpha t e^{-k_b t} = \frac{\alpha}{k_b}(1 - e^{-k_b t}),\end{aligned} \tag{72}$$

whose form it the same as Eq.(25), with k_b playing the role of the degradation rate constant β. Also note, however, that the distribution here is *not* Poissonian. The analytic formula Eq.(69) is shown in Figure 3 for initial distribution is $(v_A(n), v_I(n)) = (\delta_{n,0}, 0)$ and the parameters $(k_f/\alpha, k_b/\alpha) = (0.0, 0.1)$, for the time points $\alpha t = 3.0$ and 7.0, where the values of $P(n,t)$ are connected with solid lines. The results for $k_b = 0.0$ are also shown with dashed lines, where the other parameters and the initial distribution are the same as those of the solid lines. We see that the average number of molecules for $k_b = 0.1\alpha$ is less than that for $k_b = 0$ as expected, and the distribution develops a second peak at $n = 0$. The Gillespie simulation [35] results for $k_b = 0.1\alpha$ are also shown with filled circles, which show almost perfect agreement with the analytic formula.

As another check for the expression in Eq.(69), the results for $\alpha t = 3.0$, 5.0, and 7.0, for the initial distribution $(v_A(n), v_I(n)) = (0.5\delta_{n,0}, 0.5\delta_{n,0})$ and the parameters $(k_f/\alpha, k_b/\alpha) = (0.5, 0.1)$ are compared with the results of the Gillespie stochastic simulations [35] in Fig. 4. Again the agreement is almost perfect, indicating the validity of the expression in Eq.(69).

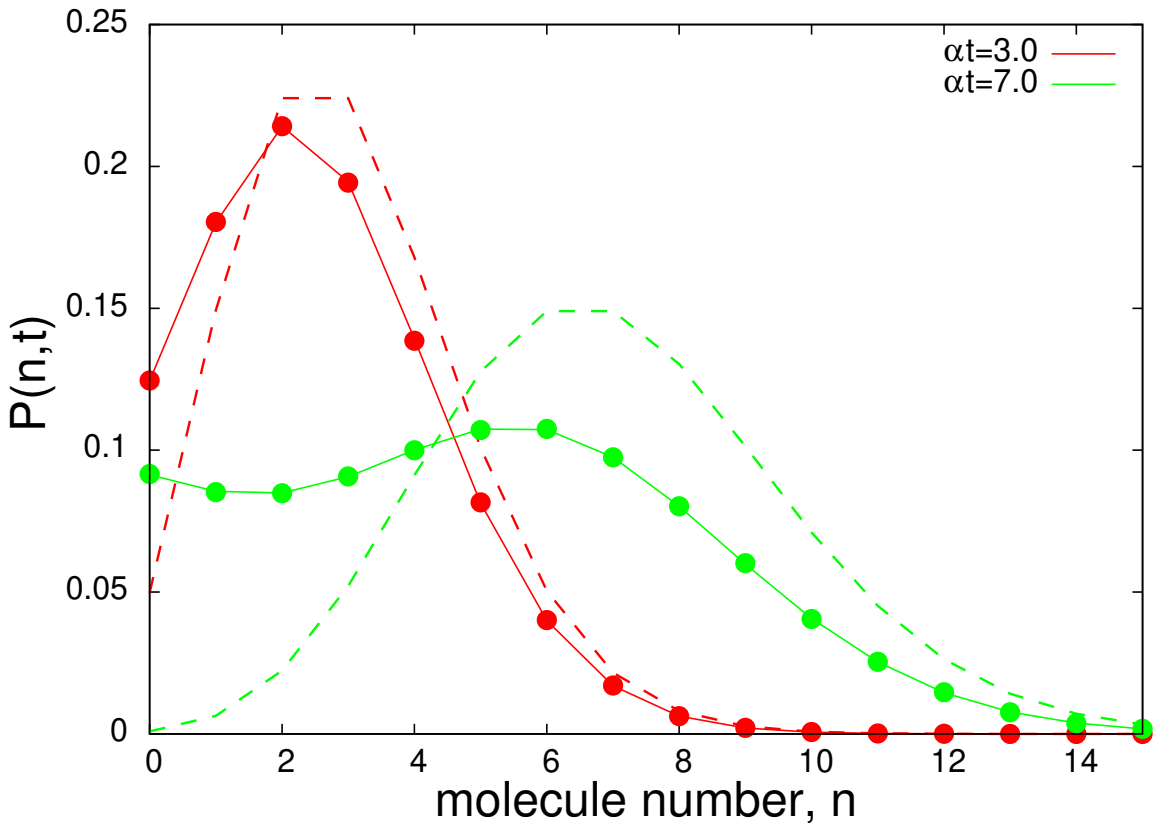

FIG. 3. The probability distribution of the molecule numbers for the model with stochastic rates for $\alpha t = 3.0$ and 7.0, obtained using the analytic formula Eq.(69) (solid lines). The initial distribution is $(v_A(n), v_I(n)) = (\delta_{n,0}, 0)$, and the parameters are $k_f = 0$ and $k_b = 0.1\alpha$. The filled circles are the results of the stochastic simulations. The results of 10^6 independent simulations were averaged. The distributions at the same time points with $k_b = 0$, obtained using analytic formula Eq.(10), are shown with dashed lines for comparison, for the same initial distribution and the parameters. The dependence on the initial molecule numbers is omitted in the vertical label for notational simplicity.

X. DISCUSSION

Poisson distributions appear ubiquitously in stochastic dynamics of gene expression, and the Poisson noise is considered to be the most basic type of noise when analyzing various

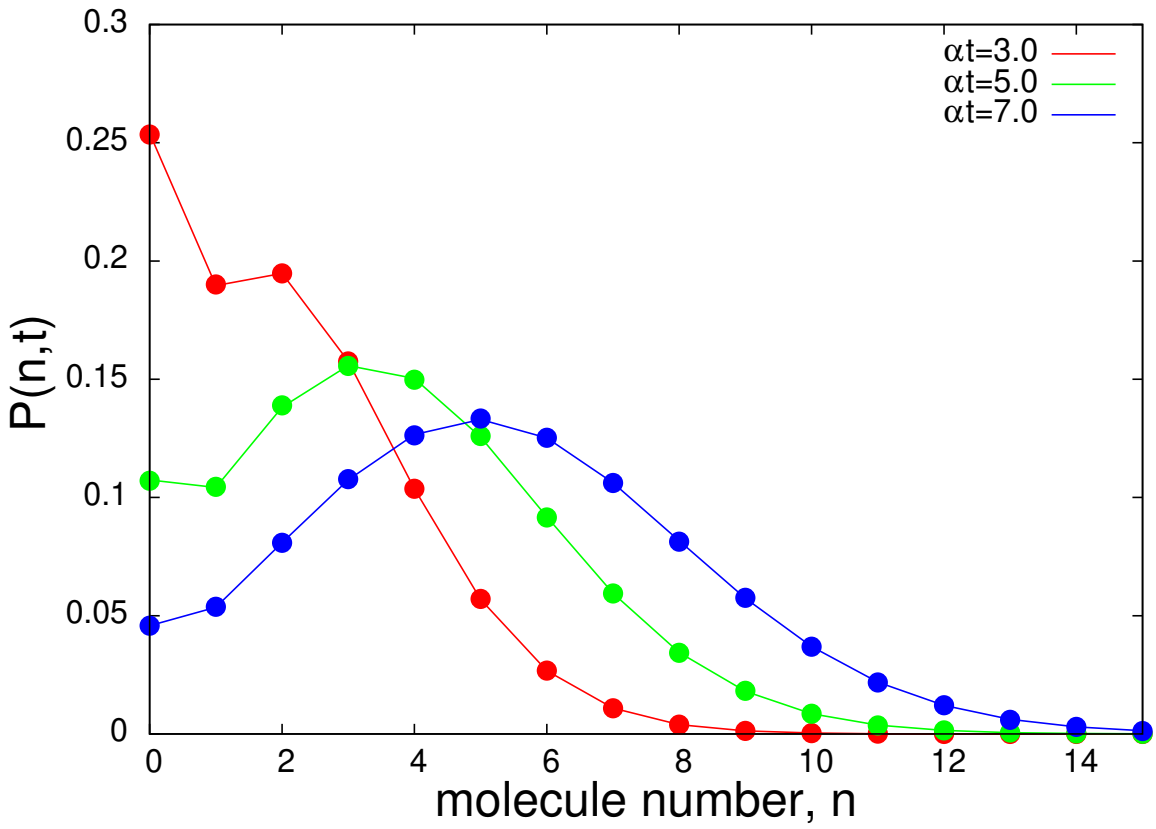

FIG. 4. The probability distribution of the molecule numbers for the model with stochastic rates for $\alpha t = 3.0$, 5.0 and 7.0, obtained using the analytic formula Eq.(69) (solid lines). The initial distribution is $(v_A(n), v_I(n)) = (0.5\delta_{n,0}, 0.5\delta_{n,0})$, and the parameters are $k_f = 0.5\alpha$ and $k_b = 0.1\alpha$. The symbols are the results of the stochastic simulations. The results of 10^6 independent simulations were averaged. The dependence on the initial molecule numbers is omitted in the vertical label for notational simplicity.

components of stochastic fluctuations. However, when gene products are allowed to get degraded, it has not been clear whether the molecule number following a Poisson distribution is equal to the number of certain independent events in time, and if this is the case, what the corresponding events are. I answered this question in this work, by showing that the number of molecules distributed according to the Poisson distribution is equal to the number

of creations of the molecules that are destined to survive until the end of a given period, which are indeed independent events in time that form an inhomogeneous Poisson process. Using this viewpoint, I could derive the Poisson distribution not only for the Markovian model with time-dependent rates but also for the model with time-delayed degradation without performing the difficult task of solving the non-Markovian master equation. I could also derive general time-dependent probability distribution in a multispecies monomolecular reaction model that allows species with non-exponential distributions of lifetimes, of which the stochastic gene dynamics with delayed degradation is a special case. By superposing these distributions with different parameters, I also derived a novel series representation for the molecule number distribution in the telegraph model without degradation.

The expansion in terms of the shifted Poisson distribution introduced in the current work, where μ is fixed and the values of n_0 vary, should be distinguished from the Poisson representation [42] where the Poisson distributions with varying μ are used in the expansion. For the model with stochastic rate, the expansion was performed with varying values of both n_0 and μ. The fact that I found a novel series solution Eq. (69) of the telegraph model, which is almost impossible to obtain from the generating function, suggests the usefulness of the current approach.

Of course, the molecule number follows the Poisson distribution only under the simplifying assumption of independent creation events. In the case of the protein, the number of protein molecules follows the Poisson distribution only if we approximate the transcription and the translation as a one-step process. In reality, the mRNA molecule has a non-zero lifetime, during which protein gets translated at a certain rate, leading to bursty translations [8–10]. The creations of protein molecules are not a Poisson process in such a model, since the translation rate depends on the mRNA concentration.

Despite these limitations, general analytic solutions found in this work may be of value as a basis for perturbations to obtain more realistic descriptions of the gene-regulatory network. Since many non-Poisson processes can be approximated by models with stochastic rates where we have a Poisson process for a given realization of the rate variable, the probability distribution can be obtained as a weighted average of the distributions obtained for Poisson processes. Even for models where one cannot perform analytic integration as was done here in the case of the telegraph model, one may numerically solve the master equation and then use the analytic solution for a given realization of the rate variable, so that the

expectation values of various physical quantities are expressed as weighted averages over numerical distributions of the stochastic rates. Similar analyses may be done for other sophisticated models of gene regulatory networks by combining the analytic solutions found in the current work with other analytic solutions and/or numerical computations.

ACKNOWLEDGMENTS

The author thanks anonymous reviewers for valuable comments and suggestions that lead to the addition of sections VII and VIII. This work was supported by the National Research Foundation of Korea, funded by the Ministry of Science and ICT (NRF-2020R1A2C1005956).

---

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

[2] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Stochastic gene expression in a single cell, Science **297**, 1183 (2002).

[3] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, Proc. Natl. Acad. Sci. **99**, 12795 (2002).

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

[5] J. M. Raser and E. K. O'Shea, Control of stochasticity in eukaryotic gene expression, Science **304**, 1811 (2004).

[6] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Real-time kinetics of gene activity in individual bacteria, Cell **123**, 1025 (2005).

[7] J. R. S. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman, Single-cell proteomic analysis of s. cerevisiae reveals the architecture of biological noise, Nature (London) **441**, 840 (2006).

[8] L. Cai, N. Friedman, and X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, Nature (London) **440**, 358 (2006).

[9] N. Friedman, L. Cai, and X. S. Xie, Linking stochastic dynamics to population distribution: An analytical framework of gene expression, Phys. Rev. Lett. **97**, 168302 (2006).

[10] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, Probing gene expression in live cells, one protein molecule at a time, Science **311**, 1600 (2006).

[11] A. Lipshtat, A. Loinger, N. Q. Balaban, and O. Biham, Genetic toggle switch without cooperative binding, Phys. Rev. Lett. **96**, 188101 (2006).

[12] A. Raj and A. van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences, Cell **135**, 216 (2008).

[13] A. Eldar and M. B. Elowitz, Functional roles for noise in genetic circuits, Nature **467**, 167 (2010).

[14] A. Ochab-Marcinek and M. Tabaka, Bimodal gene expression in noncooperative regulatory systems, Proc. Natl. Acad. Sci. **107**, 22096 (2010).

[15] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying e. coli proteome and transcriptome with single-molecule sensitivity in single cells, Science **329**, 533 (2010).

[16] J. Miękisz, J. Poleszczuk, M. Bodnar, and U. Foryś, Stochastic models of gene expression with delayed degradation, Bull. Math. Biol. **73**, 2231 (2011).

[17] L. F. Lafuerza and R. Toral, Exact solution of a stochastic protein dynamics model with delayed degradation, Phys. Rev. E **84**, 051121 (2011).

[18] G. Neuert, B. Munsky, R. Z. Tan, L. Teytelman, M. Khammash, and A. van Oudenaarden, Systematic identification of signal-activated stochastic gene regulation, Science **339**, 584 (2013).

[19] A. Sanchez and I. Golding, Genetic determinants and cellular constraints in noisy gene expression, Science **342**, 1188 (2013).

[20] N. Kumar, T. Platini, and R. V. Kulkarni, Exact distributions for stochastic gene expression models with bursting and feedback, Phys. Rev. Lett. **113**, 268105 (2014).

[21] A. Ochab-Marcinek and M. Tabaka, Transcriptional leakage versus noise: A simple mechanism of conversion between binary and graded response in autoregulated genes, Phys. Rev. E **91**, 012704 (2015).

[22] H. Ge, H. Qian, and X. S. Xie, Stochastic phenotype transition of a single cell in an intermediate region of gene state switching, Phys. Rev. Lett. **114**, 078101 (2015).

[23] Y. R. Lim, J.-H. Kim, S. J. Park, G.-S. Yang, S. Song, S.-K. Chang, N. K. Lee, and J. Sung, Quantitative understanding of probabilistic behavior of living cells operated by vibrant intra-

cellular networks, Phys. Rev. X **5**, 031014 (2015).

[24] J. Lee and J. Lee, Quantitative analysis of a transient dynamics of a gene regulatory network, Phys. Rev. E **98**, 062404 (2018).

[25] L. Ham, R. D. Brackston, and M. P. H. Stumpf, Extrinsic noise and heavy-tailed laws in gene expression, Phys. Rev. Lett. **124**, 108101 (2020).

[26] J. Lin and A. Amir, Disentangling intrinsic and extrinsic gene expression noise in growing cells, Phys. Rev. Lett. **126**, 078101 (2021).

[27] A. Terebus, F. Manuchehrfar, Y. Cao, and J. Liang, Exact probability landscapes of stochastic phenotype switching in feed-forward loops: Phase diagrams of multimodality, Front. Genet. **12**, 645640 (2021).

[28] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin Heidelberg New York, 2nd enlarged edition, 1985).

[29] T. Jahnke and W. Huisinga, Solving the chemical master equation for monomolecular reaction systems analytically, J. Math. Biol. **54**, 1 (2007).

[30] M. Thattai, Universal poisson statistics of mrnas with complex decay pathways, Biophysical Journal **110**, 301 (2016).

[31] N. G. Van Kampen, *Stochastic Processesin Physics and Chemistry* (Elsevier, Amsterdam, 2007).

[32] S. M. Ross, *Stochastic Processes* (Wiley, Hoboken, 1996).

[33] D. Stirzaker, *Stochastic Processes & Models* (Oxford, New York, 2005).

[34] D. L. Isaacson and R. W. Madsen, *Markov Chains, Theory and Applications* (Wiley, New York, 1976).

[35] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).

[36] C. Gadgila, C. H. Lee, and H. G. Othmerb, A stochastic analysis of first-order reaction networks, Bull. Math. Biol. **67**, 901 (2005).

[37] D. Cappellettia, A. P. Majumder, and C. Wiuf, The dynamics of stochastic mono-molecular reaction systems in stochastic environments, Stochastic Processes and their Applications **137**, 106 (2021).

[38] X. Cai, Exact stochastic simulation of coupled chemical reactions with delays, J. Chem. Phys. **126**, 124108 (2007).

[39] J. Peccoud and B. Ycart, Markovian modeling of gene-product synthesis, Theor. Popul. Biol. **48**, 222 (1995).

[40] S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, Stochasticity of gene products from transcriptional pulsing, Phys. Rev. E **79**, 031911 (2009).

[41] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

[42] C. W. Gardiner and S. Chaturvedi, The poisson representation. i. a new technique for chemical master equations, J. Stat. Phys. **17**, 429 (1977).

Appendix A: The Poisson distribution as the limit of the distribution for independent trials

The number of successes in N independent trials, with the success probability at i-th trial being p_i, follows the probability distribution which is a generalization of the binomial distribution,

$$\tilde{P}_{\,\mathrm{binom}}(n;\{p_1,\cdots,p_N\}) \equiv \sum_{\{i_1<i_2\cdots<i_n\}} p_{i_1}p_{i_2}\cdots p_{i_n} \prod_{k\notin\{i_1,\cdots,i_n\}} (1-p_k) \tag{A1}$$

I want to show that the Poisson distribution can be obtained from this expression in the limit of $N\to\infty$ with $\mu=\sum_{j=1}^{N} p_j$ fixed. One can derive the desired result using the generating function $F(z)\equiv\sum_n P(n)z^n$. On one hand, the generating function $\tilde{F}(z;\{p_1,\cdots,p_N\})_{\mathrm{binom}}$ for the distribution $\tilde{P}_{\,\mathrm{binom}}(n;\{p_1,\cdots,p_N\})$ is

$$\begin{aligned}
\tilde{F}(z;\{p_1,\cdots,p_N\})_{\mathrm{binom}} &\equiv \sum z^n \tilde{P}_{\,\mathrm{binom}}(n;\{p_1,\cdots,p_N\}) \\
&= \sum_{n=0}^{N} z^n \sum_{\{i_1<i_2\cdots<i_n\}} p_{i_1}p_{i_2}\cdots p_{i_n} \prod_{k\notin\{i_1,\cdots,i_n\}} (1-p_k) \\
&= \sum_{n=0}^{N} \sum_{\{i_1<i_2\cdots<i_n\}} (zp_{i_1})(zp_{i_2})\cdots(zp_{i_n}) \prod_{k\notin\{i_1,\cdots,i_n\}} (1-p_k) \\
&= \prod_{j=1}^{N} [zp_j+(1-p_j)]
\end{aligned} \tag{A2}$$

On the other hand, the generating function $F_{\mathrm{Poisson}}(z)$ for the Poisson distribution is

$$F_{\mathrm{Poisson}}(z;\mu) = \sum_n \frac{z^n\mu^n e^{-\mu}}{n!} = e^{\mu(z-1)}. \tag{A3}$$

Now, Eq.(A2) can be rewritten as

$$\tilde{F}(z;\{p_1,\cdots,p_N\})_{\rm binom} = \prod_{j=1}^{N}[1+(z-1)p_j] = \prod_{j=1}^{N}\left[e^{(z-1)p_j}+O(p_j^2)\right]$$
$$= e^{\mu(z-1)}+O(1/N). \tag{A4}$$

Therefore,

$$\lim_{N\to\infty}\tilde{F}(z;\{p_1,\cdots,p_N\})_{\rm binom} = F_{\rm Poisson}(z;\mu), \tag{A5}$$

from which we deduce

$$\lim_{N\to\infty}\tilde{P}_{\rm\ binom}(n;\{p_1,\cdots,p_N\}) = P_{\rm\ Poisson}(n;\mu). \tag{A6}$$

Since the binomial distribution is a special case of $\tilde{P}_{\rm\ binom}(n;\{p_1,\cdots,p_N\})$ with $p_j=p$ for all j, it is easy to see that the Poisson distribution is obtained from the binomial distribution by taking the limit of $N\to\infty$ with $\mu=Np$ fixed.

Appendix B: The superposition of binomial distribution weighted by a Poisson distribution is again a Poisson distribution

Consider a binomial distribution for n_0 independent and identical trials with success probability p at each trial, and suppose that n_0 is itself stochastic, distributed with Poisson distribution with the expectation value μ. Then the number n of success again follows a Poisson distribution, with the expected number of events being μp [32]:

$$\sum_{n_0} P_{\rm Poisson}(n_0;\mu)P_{\rm binom}(n;\{n_0,p\}) = P_{\rm Poisson}(n;\mu p). \tag{B1}$$

The Poisson distribution at the left-hand side of Eq.(B1) can be considered to come from a Poisson process where an event happens within a short time interval $[t,t+dt]$ with probability $\lambda(t)dt$ so that $\mu=\int_0^t\lambda(t')dt'$. Now, whenever such an event happens, we also toss a coin with head probability p and count the event only when the coin produces the head. It is intuitively clear that the number of counted events in the time interval $[0,t]$ follows the Poisson distribution with the expected number μp, as given in the right-hand side of Eq.(B1). Before formally proving Eq.(B1), I first prove the discrete version obtained by replacing the Poisson distributions in Eq.(B1) with the binomial distributions,

$$\sum_{n_0} P_{\rm binom}(n_0;\{N,q\})P_{\rm binom}(n;\{n_0,p\}) = P_{\rm binom}(n;\{N,qp\}), \tag{B2}$$

which may be easier to grasp intuitively. This expression arises in the situation where we toss two coins N times, the probability of the head for two coins at each trial being p and q, respectively. The success of a given trial is defined as the event that both coins produce heads. Then the number of successes follows the binomial distribution with the success probability pq at each trial, which is the right-hand side of Eq.(B2). The left-hand side is its decomposition using conditional probability. We first compute the probability $P_{\rm binom}(n_0;\{N,q\})$ that the first coin produced heads n_0 times. We then compute the probability $P_{\rm binom}(n;\{n_0,p\})$ that *among* n_0 trials with the first coin producing heads, n of them have the second coin producing the heads. It is intuitively clear that the summation on the left-hand side is equal to the right-hand side, but one can also explicitly show that

$$\begin{aligned}
&\sum_{n_0} P_{\rm binom}(n_0;\{N,q\})P_{\rm binom}(n;\{n_0,p\}) \\
&= \sum_{n_0} \frac{N!}{n_0!(N-n_0)!}q^{n_0}(1-q)^{N-n_0} \times \frac{n_0!}{n!(n_0-n)!}p^n(1-p)^{n_0-n} \\
&= \frac{N!}{(N-n)!n!}(pq)^n \sum_{n_0} \frac{(N-n)!}{(N-n_0)!(n_0-n)!}q^{n_0-n}(1-p)^{n_0-n}(1-q)^{N-n_0} \\
&= \frac{N!}{(N-n)!n!}(pq)^n \sum_{j} \frac{(N-n)!}{(N-n-j)!j!}[q(1-p)]^j(1-q)^{N-n-j} \\
&= \frac{N!}{(N-n)!n!}(pq)^n \left[q(1-p)+1-q\right]^{N-n} = \frac{N!}{(N-n)!n!}(pq)^n \left[1-pq\right]^{N-n} \\
&= P_{\rm binom}(n;\{N,qp\}),
\end{aligned} \tag{B3}$$

proving Eq.(B2). It is straightforward to extend the proof to the inhomogeneous case where p and q are different for each trial, and Eq.(B1) is then obtained in the limit of $N\to\infty$ with $\mu=\sum_{i=1}^N q_i$ fixed. However, one can also prove Eq.(B1) directly:

$$\begin{aligned}
&\sum_{n_0} P_{\rm Poisson}(n_0;\mu)P_{\rm binom}(n;\{n_0,p\}) \\
&= \sum_{n_0} \frac{e^{-\mu}\mu^{n_0}}{n_0!} \times \frac{n_0!}{n!(n_0-n)!}p^n(1-p)^{n_0-n} = \frac{e^{-\mu}}{n!}(\mu p)^n \sum_{n_0} \frac{\mu^{n_0-n}}{(n_0-n)!}(1-p)^{n_0-n} \\
&= \frac{e^{-\mu}}{n!}(\mu p)^n \sum_{j} \frac{[\mu(1-p)]^j}{j!} = \frac{e^{-\mu}}{n!}(\mu p)^n e^{\mu(1-p)} = \frac{e^{-\mu p}}{n!}(\mu p)^n = P_{\rm Poisson}(n;\mu p).
\end{aligned} \tag{B4}$$

Appendix C: The time-dependent distribution of the Markovian model with nonzero initial numbers of molecules (Eq.(27)) is the solution of the master equation (23).

To show that the distribution given by Eq.(27) is the solution of the master equation, it is convenient to use the generating function $F(z,t) \equiv \sum_n P(n,t)z^n$. Then the master equation (23) turns into

$$\partial_t F(z,t) = (z-1)(\alpha(t) - \beta(t)\partial_z)F(z,t). \tag{C1}$$

The generating function for the distribution in Eq.(27) is

$$\begin{aligned} F(z,t) &= \sum_n z^n \sum_{n_1} \frac{n_0!}{n_1!(n_0-n_1)!} \frac{e^{-\mu(t)}\mu(t)^{n-n_1}}{(n-n_1)!} p_{\rm surv}(t|0)^{n_1}(1-p_{\rm surv}(t|0))^{n_0-n_1} \\ &= \sum_{n_1} z^{n_1} \frac{n_0!}{n_1!(n_0-n_1)!} p_{\rm surv}(t|0)^{n_1}(1-p_{\rm surv}(t|0))^{n_0-n_1} \times \sum_{n_2} \frac{e^{-\mu(t)}\mu(t)^{n_2}z^{n_2}}{(n_2)!} \\ &= \left[1 + p_{\rm surv}(t|0)(z-1)\right]^{n_0} e^{\mu(t)(z-1)} \end{aligned} \tag{C2}$$

On one hand, by substituting $F(z)$ to the left-hand side of Eq.(C1), we get

$$\begin{aligned} \partial_t F(z,t) = {} & n_0(z-1)\left[1 + p_{\rm surv}(t|0)\right]^{n_0-1} \dot{p}_{\rm surv}(t|0) e^{\mu(t)(z-1)} \\ & + (z-1)\left[1 + p_{\rm surv}(t|0)(z-1)\right]^{n_0} e^{\mu(t)(z-1)} \dot{\mu}(t). \end{aligned} \tag{C3}$$

On the other hand, by substituting $F(z)$ to the right-hand side of Eq.(C1), we get

$$\begin{aligned} (z-1)(\alpha(t) - \beta(t)\partial_z)F(z,t) = {} & \alpha(t)(z-1)\left[1 + p_{\rm surv}(t|0)(z-1)\right]^{n_0} e^{\mu(t)(z-1)} \\ & - n_0\beta(t)(z-1)\left[1 + p_{\rm surv}(t|0)(z-1)\right]^{n_0-1} p_{\rm surv}(t|0) e^{\mu(t)(z-1)} \\ & - \beta(t)(z-1)\mu(t)\left[1 + p_{\rm surv}(t|0)(z-1)\right]^{n_0} e^{\mu(t)(z-1)} \end{aligned} \tag{C4}$$

Since

$$\begin{aligned} \dot{p}_{\rm surv}(t|0) &= -\beta(t) p_{\rm surv}(t|0), \\ \dot{\mu}(t) &= \alpha(t) - \beta(t)\mu(t), \end{aligned} \tag{C5}$$

which can be easily checked by taking the time derivatives of $p_{\rm surv}(t|0)$ and $\mu(t)$ in Eqs. (21) and (22), we see that the expressions in Eq.(C3) and Eq.(C4) are equal. Therefore, the distribution in Eq.(27) is the solution of the master equation (23).

Appendix D: The Poisson distribution is the limit of the distribution for independent trials, when there are multiple alternatives

The binomial distribution arises in independent and identical trials when there are only two alternatives at each trial. When there are multiple alternatives, whose number is m, then the numbers of outcomes $(n_1, \cdots n_{m-1})$ in the total N trials follow the multinomial distribution

$$P_{\rm mult}(n_1, \cdots, n_{m-1}; \{N, p_1, \cdots p_{m-1}\}) \equiv \frac{N!}{n_1! n_2! \cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} \tag{D1}$$

if the probability of k-the alternative happening at each trial is p_k, where $n_m = N - \sum_{i=1}^{m-1} n_i$ and $p_m = 1 - \sum_{i=1}^{m-1} p_i$. The corresponding generating function is:

$$\begin{aligned} F_{\rm mult}(z_1, z_2 \cdots z_{m-1}) &\equiv \sum_{n_1, \cdots n_{m-1}} P_{\rm mult}\left(n_1, \cdots, n_{m-1}\right) z_1^{n_1} \cdots z_{m-1}^{n_{m-1}} \\ &= \sum_{n_1, \cdots n_{m-1}} \frac{N!}{n_1! n_2! \cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} z_1^{n_1} \cdots z_{m-1}^{n_{m-1}} \\ &= \left[p_m + z_1 p_1 + \cdots + z_{m-1} p_{m-1}\right]^N \end{aligned} \tag{D2}$$

It is straightforward to write down the generating function for the inhomogeneous counterpart,

$$\begin{aligned} \tilde{F}_{\rm mult}(z_1, z_2 \cdots z_{m-1}) &= \prod_{j=1}^{N} \left[p_m^{(j)} + z_1 p_1^{(j)} + \cdots + z_{m-1} p_{m-1}^{(j)}\right] \\ &= \prod_{j=1}^{N} \left[1 + (z_1 - 1) p_1^{(j)} + \cdots + (z_{m-1} - 1) p_{m-1}^{(j)}\right], \end{aligned} \tag{D3}$$

where $p_k^{(j)}$ denotes the probability of the occurrence of k-th alternative happening at j-th trial, with $p_m^{(j)} \equiv 1 - \sum_{k=1}^{m-1} p_k^{(j)}$.

We now take the limit $N \to \infty$, with $\mu_k = \sum_j p_k^{(j)}$ fixed for $1 \le k \le m-1$. We get

$$\begin{aligned} \tilde{F}_{\rm mult}(z_1, z_2 \cdots z_{m-1}) &= \prod_{j=1}^{N} \left[\exp\left((z_1 - 1) p_1^{(j)} + \cdots + (z_{m-1} - 1) p_{m-1}^{(j)}\right) + O(1/N^2)\right] \\ &= \exp\left((z_1 - 1) \sum_{j=1}^{N} p_1^{(j)} + \cdots + (z_{m-1} - 1) \sum_{j=1}^{N} p_{m-1}^{(j)}\right) + O(1/N) \\ &\xrightarrow[N \to \infty]{} \exp\left((z_1 - 1)\mu_1 + \cdots + (z_{m-1} - 1)\mu_{m-1}\right), \end{aligned} \tag{D4}$$

which is nothing but the generating function for $m-1$ independent Poisson distributions, with expected number of occurrences of k-th alternative being μ_k.

Note that for finite N, k events with $k = 1, \cdots m-1$ are exclusive events and are therefore not independent. However, even if we assume they are independent, the probability of such events occurring more than once becomes negligible in the limit $N \to \infty$, because they are rare events with probability being $O(1/N)$. Therefore, the exclusive and independent events become indistinguishable. Let us illustrate this point with the homogeneous case with $m = 3$. The corresponding multinomial distribution describes the case where we toss a three-faced coin, with two heads denoted as A and B. There are three possible outcomes at each trial: The head A with probability p, the head B with probability q, or the tail with probability $1 - p - q$. The probability distribution is given by the multinomial distribution,

$$P(n_A, n_B) = \frac{N!}{n_A! n_B! (N - n_A - n_B)!} p^{n_A} q^{n_B} (1 - p - q)^{N - n_A - n_B}, \tag{D5}$$

which was already shown above to approach the Poisson distribution

$$P(n_A, n_B) = \frac{e^{-\mu_A - \mu_B}}{n_A! n_B!} \mu_A^{n_A} \mu_B^{n_B}. \tag{D6}$$

in the limit of $N \to \infty$ with $\mu_A = Np$ and $\mu_B = Nq$ fixed. Now compare this with the case where we toss two independent two-sided coins denoted as A and B at each trial, which produce heads with probabilities p and q, respectively. In contrast to the previous model, outcomes of head A and head B are independent, and the simultaneous heads of A and B are now allowed so that there are four outcomes at each trial. The probability distribution for n_A and n_B is now the product of binomial distributions,

$$P(n_A, n_B) = \frac{N!}{n_A! (N - n_A)!} p^{n_A} (1 - p)^{N - n_A} \frac{N!}{n_B! (N - n_B)!} q^{n_B} (1 - q)^{N - n_B}. \tag{D7}$$

Since each binomial distribution approaches the Poisson distribution, we again get Eq.(D6) in the limit of $N \to \infty$ with μ_A and μ_B fixed. In this limit, the probability pq of both coins producing heads, being $O(1/N^2)$, becomes negligible. Therefore, independent events happening in the short time interval of size $O(N^{-1})$ become effectively exclusive, or vice versa, in the limit of $N \to \infty$. This is the reason why multinomial distribution, or its inhomogeneous counterpart, factorizes into independent Poisson distributions in this limit.

Appendix E: The time-dependent distribution of the non-Markovian model with a non-zero initial number of molecules (Eq.(41)) is the solution of the master equation (31).

To show that the distribution given by Eq.(41) is the solution of the master equation, it is convenient to use the generating function $G(z,w,t) \equiv \sum_{n_A} \sum_{n_I} P(n_A, n_I, t) z^{n_A} w^{n_I}$. Then the master equation (31) turns into [17]

$$\begin{aligned}\partial_t G(z,w,t) = &\left[\gamma(1-z) + \beta(w-z)\right] \partial_z G + \zeta(1-w)\partial_w G + \alpha(z-1)G(z,w,t) \\ &+ \beta e^{-\zeta\tau}(1-w) \sum_{n_A} \left[1 + \Phi(z,w,\tau)\right]^{n_A - 1} n_A P(n_A, t-\tau) \\ &\times \exp\left[\alpha \int_0^\tau dt' \Phi(z,w,t')\right] \theta(t-\tau), \end{aligned} \tag{E1}$$

where

$$\Phi(z,w,t) \equiv (z-1)p_A(t) + (w-1)p_I(t) \tag{E2}$$

with $p_A(t)$ and $p_I(t)$ given in Eq.(39). The last term in Eq.(E1) involving $\Phi(z,w,t)$ describes the process where the active molecule that entered the degradation process at $t-\tau$ gets degraded. Therefore, for the initial condition where $n_I = 0$ at $t = 0$, this term is absent for $t < \tau$, implemented here by the step function $\theta(t-\tau)$.

From Eq.(41), we get the marginal probability distribution:

$$
\begin{aligned}
P(n_A,t) &= \sum_{n_I}\sum_{n'_A}\sum_{n'_I} \frac{n_0!}{n'_A!n'_I!(n_0-n'_A-n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\
&\quad \times (1-p_A(t)-p_I(t))^{n_0-n'_A-n'_I} \\
&\quad \times \frac{e^{-\mu_A(t)-\mu_I(t)}\mu_A(t)^{n_A-n'_A}\mu_I(t)^{n_I-n'_I}}{(n_A-n'_A)!(n_I-n'_I)!} \\
&= \sum_{n'_A}\sum_{n'_I} \frac{n_0!}{n'_A!n'_I!(n_0-n'_A-n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\
&\quad \times (1-p_A(t)-p_I(t))^{n_0-n'_A-n'_I} \\
&\quad \times \frac{e^{-\mu_A(t)-\mu_I(t)}\mu_A(t)^{n_A-n'_A}}{(n_A-n'_A)!} \sum_{n_I} \frac{\mu_I(t)^{n_I}}{n_I!} \\
&= \sum_{n'_A}\sum_{n'_I} \frac{n_0!}{n'_A!n'_I!(n_0-n'_A-n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\
&\quad \times (1-p_A(t)-p_I(t))^{n_0-n'_A-n'_I} \frac{e^{-\mu_A(t)}\mu_A(t)^{n_A-n'_A}}{(n_A-n'_A)!} \\
&= \sum_{n'_A} \frac{n_0!}{n'_A!(n_0-n'_A)!} p_A(t)^{n'_A}(1-p_A(t))^{n_0-n'_A} \frac{e^{-\mu_A(t)}\mu_A(t)^{n_A-n'_A}}{(n_A-n'_A)!} \\
&= \sum_{n'_A} P_{\text{binom}}(n'_A;\{n_0,p_A(t)\}) P_{\text{Poisson}}(n_A-n'_A;\mu_A(t)),
\end{aligned}
\tag{E3}
$$

and the generating function for $t \geq 0$:

$$
\begin{aligned}
G(z,w,t) &= \sum_{n_A}\sum_{n_I}\sum_{n'_A}\sum_{n'_I} \frac{z^{n_A}w^{n_I}n_0!}{n'_A!n'_I!(n_0-n'_A-n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\
&\quad \times (1-p_A(t)-p_I(t))^{n_0-n'_A-n'_I} \\
&\quad \times \frac{e^{-\mu_A(t)-\mu_I(t)}\mu_A(t)^{n_A-n'_A}\mu_I(t)^{n_I-n'_I}}{(n_A-n'_A)!(n_I-n'_I)!} \\
&= \sum_{n'_A}\sum_{n'_I} \frac{z^{n'_A}w^{n'_I}n_0!}{n'_A!n'_I!(n_0-n'_A-n'_I)!} p_A(t)^{n'_A} p_I(t)^{n'_I} \\
&\quad \times (1-p_A(t)-p_I(t))^{n_0-n'_A-n'_I} \\
&\quad \times \sum_{n''_A}\sum_{n''_I} z^{n''_A}w^{n''_I} \frac{e^{-\mu_A(t)-\mu_I(t)}\mu_A(t)^{n''_A}\mu_I(t)^{n''_I}}{(n''_A)!(n''_I)!} \\
&= [1+p_A(t)(z-1)+p_I(t)(w-1)]^{n_0}\, e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)} \\
&= [1+\Phi(z,w,t)]^{n_0}\, e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}.
\end{aligned}
\tag{E4}
$$

First, we compute the summation in the last term at the right-hand side of Eq.(E1) by

substituting the expressions for $P(n_A, t)$ and $G(z, w, t)$:

$$
\begin{aligned}
&\sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A - 1} n_A P(n_A, t - \tau) \\
&= \sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A - 1} n_A \sum_{n'_A} P_{\text{binom}}(n'_A; \{n_0, p_A(t-\tau)\}) P_{\text{Poisson}}(n_A - n'_A; \mu_A(t-\tau)) \\
&= \sum_{n'_A} \sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A - 1} n_A P_{\text{binom}}(n'_A; \{n_0, p_A(t-\tau)\}) P_{\text{Poisson}}(n_A - n'_A; \mu_A(t-\tau)) \\
&= \sum_{n'_A} \sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A + n'_A - 1} (n_A + n'_A) P_{\text{binom}}(n'_A; \{n_0, p_A(t-\tau)\}) P_{\text{Poisson}}(n_A; \mu_A(t-\tau)) \\
&= \sum_{n'_A} \left[1 + \Phi(z, w, \tau)\right]^{n'_A} \frac{n_0! p_A(t-\tau)^{n'_A}}{n'_A! (n_0 - n'_A)!} (1 - p_A(t-\tau))^{n_0 - n'_A} \\
&\times \sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A - 1} e^{-\mu_A(t-\tau)} \frac{\mu_A (t-\tau)^{n_A}}{(n_A - 1)!} \\
&+ \sum_{n'_A} \left[1 + \Phi(z, w, \tau)\right]^{n'_A - 1} \frac{n_0! p_A(t-\tau)^{n'_A}}{(n'_A - 1)! (n_0 - n'_A)!} (1 - p_A(t-\tau))^{n_0 - n'_A} \\
&\times \sum_{n_A} \left[1 + \Phi(z, w, \tau)\right]^{n_A} e^{-\mu_A(t-\tau)} \frac{\mu_A (t-\tau)^{n_A}}{n_A!} \\
&= \left[1 + \Phi(z, w, \tau) p_A(t-\tau)\right]^{n_0} \mu_A(t-\tau) \exp\left[\mu_A(t-\tau)\Phi(z, w, \tau)\right] \\
&+ n_0 p_A(t-\tau) \left[1 + \Phi(z, w, \tau) p_A(t-\tau)\right]^{n_0 - 1} \exp\left[\mu_A(t-\tau)\Phi(z, w, \tau)\right] \\
&= \left[1 + \Phi(z, w, t)\right]^{n_0} \mu_A(t-\tau) \exp\left[\mu_A(t-\tau)\Phi(z, w, \tau)\right] \\
&+ n_0 p_A(t-\tau) \left[1 + \Phi(z, w, t)\right]^{n_0 - 1} \exp\left[\mu_A(t-\tau)\Phi(z, w, \tau)\right]
\end{aligned}
\tag{E5}
$$

where to get the last line, we used the fact that

$$
p_A(\tau) p_A(t-\tau) = p_A(t), \quad p_I(\tau) p_A(t-\tau) = p_I(t) \tag{E6}
$$

so that

$$
\begin{aligned}
\Phi(z, w, \tau) p_A(t-\tau) &= (z-1) p_A(\tau) p_A(t-\tau) + (w-1) p_I(\tau) p_A(t-\tau) \\
&= (z-1) p_A(t) + (w-1) p_I(t) = \Phi(t).
\end{aligned}
\tag{E7}
$$

Therefore, the right-hand side of Eq.(31) becomes

$$
\begin{aligned}
&[\gamma(1-z)+\beta(w-z)]\,\partial_z G+\zeta(1-w)\partial_w G+\alpha(z-1)G(z,w,t)\\
&+\beta e^{-\zeta\tau}(1-w)\sum_{n_A}\left[1+\Phi(z,w,\tau)\right]^{n_A-1} n_A P(n_A,t-\tau)\exp\left[\alpha\int_0^\tau dt'\Phi(z,w,t')\right]\theta(t-\tau)\\
&=n_0\left[1+\Phi(z,w,t)\right]^{n_0-1}\left[(-a(z-1)+\beta(w-1))\,p_A(t)-\zeta(w-1)p_I(t)\right]e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&+\left[1+\Phi(z,w,t)\right]^{n_0}\left[(-a(z-1)+\beta(w-1))\,\mu_A(t)-\zeta(w-1)\mu_I(t)+\alpha(z-1)\right]e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&-\beta e^{-\zeta\tau}(w-1)\mu_A(t-\tau)\left[1+\Phi(z,w,t)\right]^{n_0}\exp\left[\mu_A(t-\tau)\Phi(z,w,\tau)+\alpha\int_0^\tau dt'\Phi(z,w,t')\right]\theta(t-\tau)\\
&-\beta e^{-\zeta\tau}(w-1)n_0 p_A(t-\tau)\left[1+\Phi(z,w,t)\right]^{n_0-1}\exp\left[\mu_A(t-\tau)\Phi(z,w,\tau)+\alpha\int_0^\tau dt'\Phi(z,w,t')\right]\theta(t-\tau)\\
&=n_0\left[1+\Phi(z,w,t)\right]^{n_0-1}\left[(-a(z-1)+\beta(w-1))\,p_A(t)-\zeta(w-1)p_I(t)\right]e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&+\left[1+\Phi(z,w,t)\right]^{n_0}\left[(-a(z-1)+\beta(w-1))\,\mu_A(t)-\zeta(w-1)\mu_I(t)+\alpha(z-1)\right]e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&-\beta e^{-\zeta\tau}(w-1)\mu_A(t-\tau)\left[1+\Phi(z,w,t)\right]^{n_0}e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\theta(t-\tau)\\
&-\beta e^{-\zeta\tau}(w-1)\theta(t-\tau)n_0 p_A(t-\tau)\left[1+\Phi(z,w,t)\right]^{n_0-1}e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\theta(t-\tau)
\end{aligned}
\tag{E8}
$$

where to get the last line, I used the fact that

$$
\mu_A(t-\tau)p_A(\tau)=\mu_A(t)-\mu_A(\tau),\quad \mu_A(t-\tau)p_I(\tau)=\mu_I(t)-\mu_I(\tau),
\tag{E9}
$$

and

$$
\alpha\int_0^\tau\Phi(t')dt'=(z-1)\mu_A(\tau)+(w-1)\mu_I(\tau),
\tag{E10}
$$

so that

$$
\exp\left[\mu_A(t-\tau)\Phi(z,w,\tau)+\alpha\int_0^\tau dt'\Phi(z,w,t')\right]=e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}.
\tag{E11}
$$

Next, substituting the expression for $G(z,w,t)$ in Eq.(E4) to the left-hand side of Eq.(E1), we get

$$
\begin{aligned}
\partial_t G(z,w,t)&=n_0\left[1+\Phi(z,w,t)\right]^{n_0-1}\\
&\quad\times\left[(z-1)\dot{p}_A(t)+(w-1)\dot{p}_I(t)\right]e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&\quad+\left[1+\Phi(z,w,t)\right]^{n_0}e^{\mu_A(t)(z-1)+\mu_I(t)(w-1)}\\
&\quad\times\left[(z-1)\dot{\mu}_A(t)+(w-1)\dot{\mu}_I(t)\right]
\end{aligned}
\tag{E12}
$$

Since

$$\begin{aligned}
\dot{p}_A(t) &= -ap_A(t), \\
\dot{p}_I(t) &= \beta p_A(t) - \zeta p_I(t) - \beta e^{-\zeta\tau} p_A(t-\tau)\theta(t-\tau), \\
\dot{\mu}_A(t) &= \alpha - a\ \mu_A(t), \\
\dot{\mu}_I(t) &= -\zeta\mu_I(t) + \beta\left[\mu_A(t) - e^{-\zeta\tau}\mu(t-\tau)\theta(t-\tau)\right],
\end{aligned} \tag{E13}$$

we see that the expressions in Eq.(E8) and Eq.(E12) are equal. Therefore, the distribution in Eq.(41) is the solution of the master equation (31).

Appendix F: Detailed derivation of dynamics equations for $p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})$(Eq.(48)) and $\mu_i(t)$(Eq.(50)) in a general monomolecular reaction network

To obtain $p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})$, consider its change $\Delta p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})$ after a short time interval of size Δt. Let us first define the $\ell\times\ell$ matrix $\mathbf{B}(t)$ by the components

$$\begin{aligned}
B_{ji}(t) &= c_{ij}(t) \quad (i\neq j) \\
B_{ii}(t) &= -\sum_{j=0}^{\ell} c_{ij}(t) = -\beta_i(t) - \sum_{j=1}^{\ell} c_{ij}(t).
\end{aligned} \tag{F1}$$

Some molecules of other species that was created at t' will transform to X_i during the interval $[t,t+\Delta t]$, which will contribute positively to $\Delta p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})$, which is written as

$$\Delta p_{ij}^{(1)(\mathrm{surv})}(t|t',\boldsymbol{\tau}) = \sum_{k\neq i,k\geq 1} c_{ki} p_{kj}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\Delta t = \sum_{k\neq i} B_{ik} p_{kj}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\Delta t. \tag{F2}$$

Similarly, there is a negative contribution from the X_i molecules that transform to other species or disintegrate during the interval $[t,t+\Delta t]$, written as

$$\Delta p_{ij}^{(2)(\mathrm{surv})}(t|t',\boldsymbol{\tau}) = -\sum_{k\neq i,k\geq 0} c_{ik} p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\Delta t = B_{ii} p_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\Delta t. \tag{F3}$$

For a molecule with a finite upper bound on its lifetime, we have an additional contribution if $t-\tau_i > t'$: If the molecule was converted from other species to X_i during the short interval $[t-\tau_i, t-\tau_i+\Delta t]$ and has been surviving as species i up to t without transforming to other species, then it must get degraded at t. Therefore, the corresponding contribution $\Delta p_{ij}^{(3)(\mathrm{surv})}(t|t',\boldsymbol{\tau})$ is given by the negative of the product of the probability $p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau})$ that a molecule created as X_j at t' surviving until $t-\tau_i$ as some X_k $(k\neq i)$, probability

$c_{ki}(t-\tau_i)\Delta t$ that it will transform to X_i during the short period $[t-\tau_i, t-\tau_i+\Delta t]$, and the conditional probability $\tilde{p}_i^{(survive)}(t|t-\tau_i)$ that will survive as X_i until t *without* transforming to other species,

$$\begin{aligned}\Delta p_{ij}^{(3)(\mathrm{surv})}(t|t',\boldsymbol{\tau}) &= -\Delta t \sum_{k\neq i} \tilde{p}_i^{(survive)}(t|t-\tau_i) c_{ki}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau}) \\ &= -\Delta t \sum_{k\neq i} \tilde{p}_i^{(survive)}(t|t-\tau_i) B_{ik}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau}), \end{aligned} \tag{F4}$$

where $p_{kj}^{(\mathrm{surv})}(t''|t',\boldsymbol{\tau}) \equiv 0$ for $t'' < t'$. $\tilde{p}_i^{(survive)}(t|t-\tau_i)$ is obtained by noting that for a molecule of X_i that was present at $t'' \in [t',t]$, the conditional probability that it will neither disintegrate nor transform to other species during the short period $[t''+\Delta t''] \subset [t',t]$, assuming that it does not run out of its lifetime, is given by $(1-\sum_{k=0}^{\ell} c_{ik}(t'')\Delta t'') + O((\Delta t'')^2) = (1+B_{ii}(t'')\Delta t'') + O\left((\Delta t'')^2\right)$. $\tilde{p}_i^{(survive)}(t|t-\tau_i)$ is the product of such factors, and by setting $t-\tau = N\Delta t''$ and taking the limit of $N\to\infty$, we get

$$\begin{aligned}\tilde{p}_i^{(survive)}(t|t-\tau_i) &= \lim_{N\to\infty} \mathrm{T} \prod_{k=1}^{N} \left(1 + B_{ii}\left(t-\tau_i+k\Delta t''\right)\Delta t''\right). \\ &= \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \end{aligned} \tag{F5}$$

where T in the front of the matrix indicates a time-ordered product, indicating that the matrices are multiplied from right to left in chronological order. Substituting Eq.(F5) into Eq.(F4), we get

$$\Delta p_{ij}^{(3)(\mathrm{surv})}(t|t',\boldsymbol{\tau}) = -\mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_{k\neq i} B_{ik}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau})\Delta t \tag{F6}$$

Finally, the molecule of X_i could have been created at $t-\tau_i$ and survived up to t without transforming to other species, if $t-\tau_i = t'$. The contribution from the degradation of this molecule is derived using the same logic as in the case of $\Delta p_{ij}^{(3)(\mathrm{surv})}(t|t',\boldsymbol{\tau})$,

$$\Delta p_{ij}^{(4)(\mathrm{surv})}(t|t',\boldsymbol{\tau}) = -\mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \delta(t-\tau-t')\delta_{ij}\Delta t \tag{F7}$$

Therefore, by summing Eqs.(F2), (F3), (F6), (F7), dividing by Δt and taking the limit of $\Delta t \to 0$, we get

$$\begin{aligned}\dot{p}_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau}) &= \sum_k B_{ik}(t) p_{kj}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_{k\neq i} B_{ik}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \delta_{ij}\delta(t-\tau_i-t') \end{aligned} \tag{F8}$$

Eq.(50) is then obtained by taking the derivative of Eq.(47),

$$\begin{aligned}\dot{\mu}_i(t|\boldsymbol{\tau}) &= \sum_j p_{ij}^{(\mathrm{surv})}(t|t,\boldsymbol{\tau})\alpha_j(t) + \sum_j \int_0^t \dot{p}_{ij}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\alpha_j(t')dt' \\ &= \alpha_i(t) + \sum_{j,k} B_{ik}(t) \int_0^t p_{kj}^{(\mathrm{surv})}(t|t',\boldsymbol{\tau})\alpha_j(t')dt' \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_{k\neq i}\sum_j \int_0^t B_{ik}(t-\tau_i) p_{kj}^{(\mathrm{surv})}(t-\tau_i|t',\boldsymbol{\tau})\alpha_j(t')dt' \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_j \delta_{ij}\alpha_j(t-\tau_i) \\ &= \alpha_i(t) + \sum_k B_{ik}(t)\mu_k(t|\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \sum_{k\neq i} B_{ik}(t-\tau_i)\mu_k(t-\tau_i|\boldsymbol{\tau}) \\ &\quad - \mathrm{T}\exp\left(\int_{t-\tau_i}^{t} B_{ii}(u)du\right) \alpha_i(t-\tau_i). \end{aligned} \tag{F9}$$

where $\alpha_i(t'') = \mu_i(t'') \equiv 0$ for $t'' < 0$.

Appendix G: Monomolecular reaction without finite upper bounds on molecule lifetimes

Here I explicitly show that the convolution of the multinomial distribution with Poisson distributions, derived in ref. [29], is a special case of the solution Eq.(52). This is the model in (44), where $m = \ell$ in Eq.(52). Eqs.(48) and (50) also reduce to

$$\dot{\mathbf{p}}^{(\mathrm{surv})}(t|0) = \mathbf{B}(t)\mathbf{p}^{(\mathrm{surv})}(t|0), \quad \dot{\boldsymbol{\mu}}(t) = \mathbf{B}(t)\boldsymbol{\mu}(t) + \boldsymbol{\alpha}(t), \tag{G1}$$

with their initial conditions being

$$\mathbf{p}^{(\mathrm{surv})}(0|0) = \mathbf{I}, \quad \boldsymbol{\mu}(0) = \mathbf{0}, \tag{G2}$$

where $\mathbf{p}^{(\mathrm{surv})}(t|0)$ and $\mathbf{B}(t)$ are the matrices formed by the components $p_{ij}^{(\mathrm{surv})}(t|0)$ and $B_{ij}(t)$, respectively, and $\boldsymbol{\mu}(t)$ and $\boldsymbol{\alpha}(t)$ are the column vectors formed by the components μ_i and α_i, respectively. The τ_i-dependence is omitted in the notation because they are now all fixed to ∞. Eqs.(52) along with Eqs.(G1) and (G2) are the results presented in ref. [29].

In fact, the solution of Eq.(G1) with the initial condition Eq.(G2) can be expressed in integral forms,

$$\mathbf{p}^{(\mathrm{surv})}(t|t') = \mathrm{T}\exp\int_{t'}^{t}\mathbf{B}(s)ds, \tag{G3}$$

and

$$\boldsymbol{\mu}(t) = \int_0^t \left[\mathrm{T}\exp\int_{t'}^{t}\mathbf{B}(s)ds\right]\boldsymbol{\alpha}(t')dt'. \tag{G4}$$

Eq.(G4) is the matrix generalization of Eq.(22). The integral in Eq.(G4) can be explicitly done for the case of constant rates:

$$\boldsymbol{\mu}(t) = \int_0^t e^{(t-t')\mathbf{B}}\boldsymbol{\alpha}dt' = (e^{t\mathbf{B}} - \mathbf{I})\mathbf{B}^{-1}\boldsymbol{\alpha}, \tag{G5}$$

which is a matrix generalization of Eq.(25), where $\mathbf{I}$ is the identity matrix. When all the eigenvalues of the matrix $\mathbf{B}$ have negative real parts, we get

$$\boldsymbol{\mu}(\infty) = -\mathbf{B}^{-1}\boldsymbol{\alpha}, \tag{G6}$$

which is the matrix generalization of Eq.(26).

Appendix H: The derivation of the series representation (Eq.(69)) of the general time-dependent distribution of the telegraph model.

Let us first consider the case where the gene is initially inactive, and the initial number of RNA molecules is zero. Using the series representation of the confluent hypergeometric function,

$$\Phi(a,b;z) = \sum_{\ell}\frac{(a+\ell-1)!(b-1)!}{(b+\ell-1)!(a-1)!}z^{\ell}, \tag{H1}$$

the expression in Eq.(69) for $v_I(n_0) = \delta_{n_0,0}$ and $v_A(n_0) = 0$ becomes

$$
\begin{aligned}
P(n,t) &= \sum_m \left[\frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^m (n+m)!}{n!(n+2m+1)!m!} \Phi(n+m+1, n+2m+2; (k_f - k_b - \alpha)t) \right. \\
&\left. + \frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^{m+1} (n+m)!}{n!(n+2m+2)!m!} \Phi(n+m+1, n+2m+3; (k_f - k_b - \alpha) \right] + \delta_{n,0} e^{-k_f t} \\
&= \sum_{\ell,m} \left[\frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^m \cancel{(n+m)!} \cancel{(n+2m+1)!} (n+m+\ell)!}{n! \cancel{(n+2m+1)!} m! (n+2m+\ell+1)! \cancel{(n+m)!} \ell!} ((k_f - k_b - \alpha)t)^\ell \right. \\
&\left. + \frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^{m+1} \cancel{(n+m)!} \cancel{(n+2m+2)!} (n+m+\ell)!}{n! \cancel{(n+2m+2)!} m! (n+2m+2+\ell)! \cancel{(n+m)!} \ell!} ((k_f - k_b - \alpha)t)^\ell \right] + \delta_{n,0} e^{-k_f t} \\
&= \sum_{m,\ell_1,\ell_2,\ell_3} \left[\frac{e^{-k_f t}(\alpha t)^{n+\ell_3} (k_f t)^{m+\ell_1+1} (k_b t)^{m+\ell_2} (n+m+\ell_1+\ell_2+\ell_3)! (-1)^{\ell_2+\ell_3}}{n!m!(n+2m+\ell_1+\ell_2+\ell_3+1)!\ell_1!\ell_2!\ell_3!} \right. \\
&\left. + \frac{e^{-k_f t}(\alpha t)^{n+\ell_3} (k_f t)^{m+\ell_1+1+1} (k_b t)^{m+\ell_2} (n+m+\ell_1+\ell_2+\ell_3-1)! (-1)^{\ell_2+\ell_3-1}}{n!m!(n+2m+\ell_1+\ell_2+\ell_3+1)!\ell_1!(\ell_2-1)!\ell_3!} \right] + \delta_{n,0} e^{-k_f t} \\
&= \sum_{m,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3} (k_f t)^{m+\ell_1+1} (k_b t)^{m+\ell_2} (n+m+\ell_1+\ell_2+\ell_3-1)! (-1)^{\ell_2+\ell_3}}{n!m!(n+2m+\ell_1+\ell_2+\ell_3+1)!\ell_1!\ell_2!\ell_3!} \\
&\times (n+m+\ell_1+\cancel{\ell_2}+\ell_3-\cancel{\ell_2}) + \delta_{n,0} e^{-k_f t} \\
&= \sum_{m,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3} (k_f t)^{\ell_1+1} (k_b t)^{\ell_2} (\ell_1+\ell_2+\ell_3+n-m-1)! (-1)^{\ell_2+\ell_3-m} (\ell_1+\ell_3+n)}{n!m!(\ell_1+\ell_2+\ell_3+n+1)!(\ell_1-m)!(\ell_2-m)!\ell_3!} \\
&+ \delta_{n,0} e^{-k_f t},
\end{aligned}
\tag{H2}
$$

where the shift of the summation indices $\ell_1 \to \ell_1 - m$ and $\ell_2 \to \ell_2 - m$ was performed to get the final expression. We then use the useful identity [3]

$$
\sum_m \frac{(-1)^m (p+q+r-m)!}{m!(p-m)!(q-m)!} = \frac{(p+r)!(q+r)!}{p!q!r!} \tag{H3}
$$

to perform the summation over m in Eq.(H2), to obtain

$$
P(n,t) = \sum_{\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3} (k_f t)^{\ell_1+1} (k_b t)^{\ell_2} (-1)^{\ell_2+\ell_3} (\ell_2+\ell_3+n-1)! (\ell_1+\ell_3+n)!}{n!(\ell_1+\ell_2+\ell_3+n+1)!\ell_1!\ell_2!(\ell_3+n-1)!\ell_3!} + \delta_{n,0} e^{-k_f t} \tag{H4}
$$

The expression for arbitrary value of n_0 is obtained from Eq.(H4) by making the shift $n \to n - n_0$.

[3] This identity can be checked using MATHEMATICA and be proved using mathematical induction.

Similarly, for $v_I(n_0) = 0$ and $v_A(n_0) = \delta_{n_0,0}$, the expression in Eq.(69) becomes

$$
\begin{aligned}
P(n,t) &= \sum_m \frac{e^{-k_f t}(\alpha t)^n (k_f t)^m (k_b t)^{m+1}(n+m)!}{n!(n+2m+1)!m!}\Phi(n+m+1, n+2m+2; (k_f-k_b-\alpha)t) \\
&\quad + \sum_{m\geq 0} \frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^{m+1}(n+m+1)!}{n!(n+2m+2)!(m+1)!}\Phi(n+m+2, n+2m+3; (k_f-k_b-\alpha)\Big] \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{\ell,m} \frac{e^{-k_f t}(\alpha t)^n (k_f t)^m (k_b t)^{m+1}\cancel{(n+m)!}\cancel{(n+2m+1)!}(n+m+\ell)!}{n!\cancel{(n+2m+1)!}m!(n+2m+\ell+1)!\cancel{(n+m)!}\ell!}((k_f-k_b-\alpha)t)^\ell \\
&\quad + \sum_{\ell,m\geq 0} \frac{e^{-k_f t}(\alpha t)^n (k_f t)^{m+1} (k_b t)^{m+1}\cancel{(n+m+1)!}\cancel{(n+2m+2)!}(n+m+\ell+1)!}{n!\cancel{(n+2m+2)!}(m+1)!(n+2m+2+\ell)!\cancel{(n+m+1)!}\ell!}((k_f-k_b-\alpha)t)^\ell \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{m,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{m+\ell_1}(k_b t)^{m+\ell_2+1}(n+m+\ell_1+\ell_2+\ell_3)!(-1)^{\ell_2+\ell_3}}{n!m!(n+2m+\ell_1+\ell_2+\ell_3+1)!\ell_1!\ell_2!\ell_3!} \\
&\quad + \sum_{m\geq 0,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{m+\ell_1}(k_b t)^{m+\ell_2+1}(n+m+\ell_1+\ell_2+\ell_3)!(-1)^{\ell_2+\ell_3}}{n!(m+1)!(n+2m+\ell_1+\ell_2+\ell_3+1)!(\ell_1-1)!\ell_2!\ell_3!} \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{m\geq 0,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{m+\ell_1}(k_b t)^{m+\ell_2+1}(n+m+\ell_1+\ell_2+\ell_3)!(-1)^{\ell_2+\ell_3}}{n!(m+1)!(n+2m+\ell_1+\ell_2+\ell_3+1)!\ell_1!\ell_2!\ell_3!}\times(m+1+\ell_1) \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{m\geq 0,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{\ell_1}(k_b t)^{\ell_2+1}(\ell_1+\ell_2+\ell_3+n-m)!(-1)^{\ell_2+\ell_3-m}(\ell_1+1)}{n!(m+1)!(\ell_1+\ell_2+\ell_3+n+1)!(\ell_1-m)!(\ell_2-m)!\ell_3!} \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{m,\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{\ell_1}(k_b t)^{\ell_2+1}(\ell_1+\ell_2+\ell_3+n-m+1)!(-1)^{\ell_2+\ell_3-m+1}(\ell_1+1)}{n!m!(\ell_1+\ell_2+\ell_3+n+1)!(\ell_1-m+1)!(\ell_2-m+1)!\ell_3!} \\
&\quad - \sum_{\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{\ell_1}(k_b t)^{\ell_2+1}\cancel{(\ell_1+\ell_2+\ell_3+n+1)!}(-1)^{\ell_2+\ell_3+1}\cancel{(\ell_1+1)}}{n!\cancel{(\ell_1+\ell_2+\ell_3+n+1)!}\cancel{(\ell_1+1)}\ell_1!(\ell_2+1)!\ell_3!} \\
&\quad + \frac{(\alpha t)^n e^{-(k_b+\alpha)t}}{n!} \\
&= \sum_{\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{\ell_1}(k_b t)^{\ell_2+1}(-1)^{\ell_2+\ell_3+1}(n+\ell_1+\ell_3)!(n+\ell_2+\ell_3)!}{n!(n+\ell_1+\ell_2+\ell_3+1)!(n+\ell_3-1)!\ell_1!(\ell_2+1)!\ell_3!} + \frac{(\alpha t)^n e^{-\alpha t}}{n!} \\
&= \sum_{\ell_1,\ell_2,\ell_3} \frac{e^{-k_f t}(\alpha t)^{n+\ell_3}(k_f t)^{\ell_1}(k_b t)^{\ell_2}(-1)^{\ell_2+\ell_3}(n+\ell_1+\ell_3)!(n+\ell_2+\ell_3-1)!}{n!(n+\ell_1+\ell_2+\ell_3)!(n+\ell_3-1)!\ell_1!\ell_2!\ell_3!},
\end{aligned}
\tag{H5}
$$

where I used the identity Eq.(H3) again to perform the m summation in Eq.(H5). Again, the expression for an arbitrary value of n_0 is obtained from Eq.(H5) by making the shift $n \to n - n_0$. The final expression in the Eq.(69) is obtained by taking the linear combination of Eq.(H4) and (H5) with $v_I(n_0)$ and $v_A(n_0)$ after making the shift of $n \to n - n_0$.