# Appendix to "A Logic for Checking the Probabilistic Steady-State Properties of Reaction Networks"

Vincent Picard[1,2] Anne Siegel[1]
Jérémie Bourdon[2,⋆]
[1] IRISA, CNRS, Université Rennes 1, Rennes, 35000, France
[2] LS2N, Université de Nantes, Nantes, 44322, France
jeremie.bourdon@univ-nantes.fr

# Appendix

## 0.1 Complete proof of the central limit theorem

In this appendix, we provide a detailed proof of the central limit theorem for the continuous random process.

The proof splits in two parts. First, we provide a discretization $\vec{x}(k)$ $(k \in \mathbb{N})$ of the stochastic dynamics $\vec{y}(t)$ $(t \in \mathbb{R}^+)$ that relies on a fixed time step $\delta t$. This allows to get back to simpler discrete-time Markov chain processes. We prove that the laws of $\vec{y}(k \cdot \delta t)$ and $\vec{x}(k)$ coincide when $\delta t$ tends to $0$. Secondly, we provide the Bernoulli approximation $\vec{z}(k)$ $(k \in \mathbb{N})$ of the stationary regime of $\vec{x}(k)$. Next, analytical expressions for the moments of $\vec{z}(k)$ can be derived and a central limit theorem result can be stated. Finally, we discuss the quality of the approximation for the first and second moments leading to such expressions

$$\mathbb{E}\vec{x}(k) = \mathbb{E}\vec{z}(k) + \text{error terms} \tag{1}$$

$$\text{Cov}\, \vec{x}(k) = \text{Cov}\, \vec{z}(k) + \text{error terms} \tag{2}$$

where the errors terms can be neglected in some cases.

**Step 1: discretization of the stochastic dynamics** Several discretizations of the stochastic dynamics have been proposed in order to get efficient Gillespie simulations. We introduce a similar discretization as in [**?**], with a fixed time step $\delta t$. This discretization relies on the uniformization principle of continuous Markov chains and does not require any supplementary property on the transitions probabilities of the underlying continuous Markov chain.

For the sake of simplicity, we assume that the reaction network contain the nul reaction $R_0 : \varnothing \to \varnothing$. This implies that zero or one reaction can occur at each time step $\delta t$. Let us introduce the following probability functions, including the probability of the nul function.

$$p_j(\vec{y}) = \begin{cases} \exp\left(-h_0(\vec{y})\,\delta t\right) & \text{if } j = 0, \\ \left(1 - p_0(\vec{y})\right) \frac{h_j(\vec{y})}{h_0(\vec{y})} & \text{otherwise.} \end{cases} \tag{3}$$

**Definition 1** (Discretized process)**.** *The* discretized process *of the continuous stochastic process $(\vec{y}(t))_{t\in\mathbb{R}^+}$ over the state space $\Sigma$ is the homogeneous Markov chain $(\vec{x}(k))_{k\in\mathbb{N}}$ defined over the same state space $\Sigma$, with the same initial state $\vec{x}(0) = \vec{y}(0)$ and such that the transition probabilities are defined as followed.*

$$\forall j = 0, \ldots, m, \forall \vec{y}, \vec{y}+\vec{\nu}_j \in \Sigma, \quad \mathbb{P}\left[\vec{x}(k+1) = \vec{y} + \vec{\nu}_j \mid \vec{x}(k) = \vec{y}\right] = p_j(\vec{y}). \tag{4}$$

This discretization is based upon the following principle. We consider that either one or zero reaction occurred during the time step. The probability $p_0$ for no reaction to occur is computed from the Markov chain $\vec{y}(t)$. The probabilities for each reaction to occur are proportional to those of $\vec{y}(t)$. The validity of the approximation when the time step $\delta t$ derives naturally.

**Lemma 1.** *Let $(\vec{y}'(t))$ $(t \in \mathbb{R}^+)$ be the extension to $\mathbb{R}^+$ of the discretized process $(\vec{x}(k))$ $(k \in \mathbb{N})$, such that $\forall k \in \mathbb{N}, \vec{y}'(k\delta t) = \vec{x}(k)$ and $\vec{y}'(t)$ is constant on each interval $[k\delta t, (k+1)\delta t[$. Then, for all $t \in \mathbb{R}^+$, $\vec{y}'(t)$ converges in law to $\vec{y}(t)$ when $\delta t \to 0$,*

$$\vec{y}'(t) \xrightarrow{\mathcal{L}} \vec{y}(t). \tag{5}$$

*Proof.* The process $\vec{y}'(t)$ can be defined with a random algorithm with parameters ($\tau'$,$\mu'$).

For all $t_0 \in \delta t\mathbb{N}$, let us denote $\vec{\chi} = \vec{y}'(t_0)$ the current step. Let $\tau' = \inf\{t' > 0 | \vec{y}'(t_0 + t') \neq \vec{\chi}\}$ denote the time duration until the next non-zero reaction and $\mu'$ denote the reaction used at time $(t_0 + \tau')$. In addition, let $\tau$, $\mu$ denote the time duration and the next non-zero reaction of the Gillespie algorithm.

then

$$\forall t \geq 0, \mathbb{P}[\tau' > t] \to_{\delta t \to 0} \mathbb{P}[\tau > t],$$
$$\mathbb{P}[\mu' = j] = \mathbb{P}[\mu = j].$$

Since $\vec{y}'$ changes its values at discrete timepoints $\delta t\mathbb{N}$, we have that $\tau' \in \delta t\mathbb{N}$. Therefore, $\forall t \geq 0$, $\mathbb{P}[\tau' > t] = \mathbb{P}[\tau' > \lfloor t/\delta t \rfloor \delta t] = \exp(-h_0(\vec{\chi})\delta t)^{\lfloor t/\delta t \rfloor} = \exp(-h_0(\vec{\chi})\delta t \lfloor t/\delta t \rfloor)$. This implies:

$$\exp(-h_0(\vec{\chi})\delta t(t/\delta t + 1)) \leq \mathbb{P}[\tau' > t] \leq \exp(-h_0(\vec{\chi})\delta t(t/\delta t)).$$

Therefore $\lim_{\delta t \to 0} \mathbb{P}[\tau' > t] = \exp(-h_0(\vec{\chi})t) = \mathbb{P}[\tau > t]$.

In addition, we have that

$$\begin{aligned}
\mathbb{P}[\mu' = j] &= \sum_{l \in \mathbb{N}^*} \mathbb{P}[\mu' = j, \tau' = l\delta t] \\
&= \sum_{l \in \mathbb{N}^*} \mathbb{P}[\tau' = l\delta t] \cdot \mathbb{P}[\mu' = j | \tau' = l\delta t] \\
&= \sum_{l \in \mathbb{N}^*} p_0(\vec{\chi})^{l-1}(1 - p_0(\vec{\chi})) \cdot \frac{h_j(\vec{\chi})}{h_0(\vec{\chi})} \\
&= \frac{1}{1 - p_0(\vec{\chi})}(1 - p_0(\vec{\chi}))\frac{h_j(\vec{\chi})}{h_0(\vec{\chi})} \\
&= h_j(\vec{\chi})/h_0(\vec{\chi}) = \mathbb{P}[\mu = j].
\end{aligned}$$

Therefore, the law of ($\tau'$,$\mu'$) tends to the law of ($\tau$, $\mu$) from the Gillespie algorithm. □

**Corollary 1.** *For all $k \in \mathbb{N}$, the discrete process $\vec{x}(k)$ converges in law towards $\vec{y}(k\delta t)$ when $\delta t \to 0$.*

**Step 2: Bernoulli dynamics** Let $\vec{p}$ be a probability vector with dimension $n$. The *Bernoulli variable* with parameter $\vec{p}$ is the random variable $X$ with values in $\{1, \ldots, n\}$ with the law $\mathbb{B}(\vec{p})$ defined by $\mathbb{P}[X = k] = p_k$

**Definition 2** (Bernoulli dynamics)**.** *Let $(n, m, \alpha, \beta)$ be a reaction network, $S = (\vec{\nu}_j)_{1\leq j\leq m}$ a stoichiometric matrix, $\vec{z}(0) \in \mathbb{N}^n$ is the initial state and $\vec{p}$ a probability vector with size $n$. The* Bernoulli dynamics *is the discrete random process $\vec{z}_k \in \mathbb{Z}^n$ defined by*

$$\vec{z}(k) = \vec{z}(0) + \sum_{i=1}^{k}\sum_{j=1}^{m} \delta_j^{\mu_i}\vec{\nu}_j \qquad (k \in \mathbb{N}), \tag{6}$$

*where $\mu_i$ are independent and identically distributed with law $\mathbb{B}(\vec{p})$, and $\delta_a^b$ is the Kronecker symbol. Let $\mathbb{B}(\vec{z}(0), \vec{p})$ denote the law of the process $(z(k))$ $(k \in \mathbb{N})$.*

An equivalent definition is as follows:

$$\vec{z}(k) = \vec{z}(0) + \sum_{i=1}^{k} S\vec{e}_{\mu_i}, \tag{7}$$

where $(\vec{e}_i)_{1\leq i\leq m}$ correspond to the canonical basis of $\mathbb{R}^m$.

**Proposition 1.** *Let $\vec{z}(k)$ be a Bernoulli dynamics. Then*

$$\mathbb{E}\vec{z}(k) = \vec{z}(0) + kS\vec{p}, \tag{8}$$

$$\mathrm{Cov}\ \vec{z}(k) = kS(\mathrm{diag}\,\vec{p} - \vec{p}\vec{p}^t)S^t. \tag{9}$$

*Proof.* From the linearity of expectancy, we have that $\mathbb{E}\vec{z}(k) = \vec{z}(0)+\sum_{i=1}^{k}\sum_{j=1}^{m}\mathbb{E}(\delta_j^{\mu_i}\vec{\nu}_j)$. In addition, the following holds: $\mathbb{E}(\delta_j^{\mu_i}\vec{\nu}_j) = \sum_{l=1}^{m}\mathbb{P}[\mu_i = l]\delta_j^l\vec{\nu}_j = p_j\vec{\nu}_j$. Since $S\vec{p} = \sum_{j=1}^{m} p_j\vec{\nu}_j$ we derive that $\mathbb{E}\vec{z}(k) = \vec{z}(0) + \sum_{i=1}^{k} S\vec{p} = \vec{z}(0) + kS\vec{p}$.

As co-variances are invariant under translation, we may assume that $\vec{z}(0) = \vec{0}$. From (7), we get

$$\begin{aligned}
\mathrm{Cov}(\vec{z}(k)) &= \mathbb{E}(\vec{z}(k)\vec{z}^t(k)) - \mathbb{E}\vec{z}(k)\mathbb{E}\vec{z}^t(k) \\
&= \mathbb{E}\left((\sum_{a=1}^{k} S\vec{e}_{\mu_a})(\sum_{b=1}^{k} S\vec{e}_{\mu_b})^t\right) - k^2 S\vec{p}\vec{p}^t S^t \\
&= S\left(\sum_{a=1}^{k}\sum_{b=1}^{k}\mathbb{E}(\vec{e}_{\mu_a}\vec{e}_{\mu_b}^t) - k^2\vec{p}\vec{p}^t\right)S^t.
\end{aligned}$$

When $a \neq b$, $\mu_a$ and $\mu_b$ are independent, so that $\mathbb{E}(\vec{e}_{\mu_a}\vec{e}_{\mu_b}^t) = \mathbb{E}\vec{e}_{\mu_a}\mathbb{E}\vec{e}_{\mu_b}^t = \vec{p}\vec{p}^t$. Otherwise, let $E_{i,j}$ denote the matrix with coefficient 1 at coordinates $(i,j)$ and 0 otherwise. Then, $\mathbb{E}(\vec{e}_{\mu_a}\vec{e}_{\mu_a}^t) = \sum_{l=1}^m \mathbb{P}[\mu_a = l]E_{l,l} = \operatorname{diag}\vec{p}$.

This implies that$\operatorname{Cov}(\vec{z}(k)) = S(k\operatorname{diag}\vec{p}+k(k-1)\vec{p}\vec{p}^t-k^2\vec{p}\vec{p}^t)S^t = kS(\operatorname{diag}\vec{p}-\vec{p}\vec{p}^t)S^t$. □

In addition to the analytic formulas from the moments of order 1 and 2 of the Bernoulli dynamics, we have the following asymptotic result.

**Lemma 2** (Central Limit Theorem)**.** *The Bernoulli dynamics converges in law towards a normal law:*

$$\frac{1}{\sqrt{k}}\left(\vec{z}(k) - (\vec{z}(0) + kS\vec{p})\right) \xrightarrow[k\to+\infty]{\mathcal{L}} \mathcal{N}\left(\vec{0}, W(S,\vec{p})\right),$$

*where* $W(S,\vec{p}) = S\left(\operatorname{diag}(\vec{p}) - \vec{p}\vec{p}^t\right)S^t$.

*Proof.* The central limit theorem generalized to any dimension [**?**] combined within (7) yields that

$$\begin{aligned}\frac{1}{\sqrt{k}}\left(\vec{z}(k) - (\vec{z}(0) + kS\vec{p})\right) &= \frac{1}{\sqrt{k}}\left(\sum_{i=1}^k S\vec{e}_{\mu_i} - kS\vec{p}\right)\\ &= \frac{1}{\sqrt{k}}\sum_{i=1}^k(\vec{y}_i - \vec{m})\\ &\xrightarrow[k\to+\infty]{\mathcal{L}} \mathcal{N}(\vec{0}, V)\end{aligned}$$

where $V = \operatorname{Cov} S\vec{e}_{\mu_i}$. This co-variance matrix can be easily computed as follows:

$$\begin{aligned}V &= \mathbb{E}((S\vec{e}_{\mu_i})(S\vec{e}_{\mu_i})^t) - \mathbb{E}(S\vec{e}_{\mu_i})\mathbb{E}(S\vec{e}_{\mu_i})^t\\ &= S\left(\mathbb{E}(\vec{e}_{\mu_i}\vec{e}_{\mu_i}^t)\right)S^t - S\vec{p}\vec{p}^tS^t\\ &= S\left(\operatorname{diag}\vec{p} - \vec{p}\vec{p}^t\right)S^t.\end{aligned}$$

□

Therefore, the asymptotic distribution of $\vec{z}(k)$ is a Gaussian distribution whose parameters are analytically derived from $S$ and $\vec{p}$. Intuitively, this suggest that when $k$ tends to $\infty$, $\vec{z}(k)$ is approximatively distributed according to the normal law:

$$\mathcal{N}\left(\vec{z}(0) + kS\vec{p},\ kS(\operatorname{diag}\vec{p} - \vec{p}\vec{p}^t)S^t\right). \tag{10}$$

**Step 3. Choice of the Bernoulli dynamics parameter** In order to choose a relevant parameter (probability vector $\vec{p}$) for the Bernoulli dynamics, we use a stationary-based approached. This can be viewed as an analog of the steady-state approximation of the dynamics of a differential system introduced in the Flux Balance Analysis paradigm.

**Definition 3** (Stationnary reactions probabilities)**.** *Let $\vec{x}(k)$ be the discretized process of a stochastic dynamics of a reaction network, with probability function $\vec{p}(\cdot)$. Let $\vec{x}^*(k)$ be the orthogonal projection of $\vec{x}(k)$. Assume that $\mathbb{E}\vec{p}(\vec{x}^*(k))$ has a limit when $k \to +\infty$. Then the* stationary reactions probabilities *are defined by* $\vec{\mathfrak{p}} = \lim_{k\to+\infty} \mathbb{E}\vec{p}(\vec{x}^*(k))$.

From now, we assume that we handle a Bernoulli dynamics with parameter $\vec{p} := \vec{\mathfrak{p}}$.

**Proposition 2.** *The error term between the expectancies of the discretized dynamics $\vec{x}(k)$ and the Bernoulli dynamics with parameter $\vec{\mathfrak{p}}$ is:*

$$\mathbb{E}\vec{x}(k) = \mathbb{E}\vec{z}(k) + S\sum_{l=1}^{k} \left(\mathbb{E}(\vec{p}(\vec{x}(l-1)) - \vec{\mathfrak{p}}\right).$$

*Proof.* Assume that $(\mu_i)_{i\in\mathbb{N}}$ denote the indexes of reactions hold in the simulation algorithm of $\vec{x}(k)$. Then, for all $k \in \mathbb{N}$, we have that $\vec{x}(k+1) = \vec{x}(k) + S\vec{e}_{\mu_{k+1}}$.

Therefore,

$$\begin{aligned}\mathbb{E}\vec{x}(k+1) &= \mathbb{E}\mathbb{E}(\vec{x}(k+1)|\vec{x}(k)) = \mathbb{E}\mathbb{E}(\vec{x}(k) + S\vec{e}_{\mu_{k+1}}|\vec{x}(k)) \\ &= \mathbb{E}(\vec{x}(k) + S\mathbb{E}(\vec{e}_{\mu_{k+1}}|\vec{x}(k))) = \mathbb{E}(\vec{x}(k) + S\vec{p}(\vec{x}(k))) \\ &= \mathbb{E}\vec{x}(k) + S\mathbb{E}\vec{p}(\vec{x}(k)),\end{aligned}$$

This implies that $\mathbb{E}\vec{x}(k) = \vec{x}(0) + \sum_{l=1}^{k} S\mathbb{E}\vec{p}(\vec{x}(l-1))$.

From 7 we conclude that:

$$\mathbb{E}\vec{z}(k) = \vec{z}(0) + \sum_{l=1}^{k} S\mathbb{E}\vec{e}_{\mu_l} = \vec{z}(0) + \sum_{l=1}^{k} S\vec{\mathfrak{p}}.$$

□

Let $\|A\| = \max_{1\le i\le n} \sum_{j=1}^{m} |a_{i,j}|$ denote the sup norm over $\mathbb{R}^m$.

**Corollary 2.** $\forall \varepsilon > 0$, $\exists K > 0$, $\forall k \ge K$

$$\|\mathbb{E}\vec{x}(k) - \mathbb{E}\vec{z}(k)\| \le \|S\| \sum_{l=1}^{K} \left\|\mathbb{E}(\vec{p}(\vec{x}(i-1))) - \vec{\mathfrak{p}}\right\| + (k-K)\varepsilon\|S\|. \qquad (11)$$

*Proof.* Since $\mathbb{E}\vec{x}(k) \to \vec{\mathfrak{p}}$, for every $\varepsilon > 0$, there exists $K > 0$ such that for all $k \geq K$, we have $\|\mathbb{E}\vec{x}(k) - \vec{\mathfrak{p}}\| < \varepsilon$. □

Therefore, the error term can be bounded by an affine function with slope as small as wanted.

**Step 4. Initiating the process close to equilibria points and impact of thermodynamics limit** Since the error term mainly depend on the distance between $\mathbb{E}\vec{x}(k)$ and $\vec{\mathfrak{p}}$ at the first time steps, we introduce the concept of *equilibria points*.

**Definition 4** (distance to equilibrium)**.** *Let $\vec{y} \in \mathbb{N}^n$. The* distance to equilibrium *of $\vec{y}$ is $d_e(\vec{y}) = \|\vec{p}(\vec{y}) - \vec{\mathfrak{p}}\|$. When $d_e(\vec{y}) = 0$, that is, c $\vec{p}(\vec{y}) = \vec{\mathfrak{p}}$, then $\vec{y}$ is called an* equilibrium point*.*

Let us prove that when choosing an equilibrium point as an initial point of the dynamics, then the Bernoulli approximation is exact at the thermodynamics limit. Studying a system at its thermodynamics limit consists in increasing its volume up to $\infty$ while preserving matter concentrations. A $\lambda$-*expansion* of a system, for $\lambda \in \mathbb{N}$, correspond to a system where initial matter quantities as well as the volume $\Omega$ are multiplied by $\lambda$. Performing a $\lambda$-expansion impacts as follows on the stochastic constants $c_j$, the stochastic process $\vec{y}(t)$, its discretized dynamics $\vec{x}(k)$, its Bernoulli dynamics $\vec{z}(k)$ and its Bernoulli dynamics parameter $\vec{\mathfrak{p}}$.

| Reaction network $X$ | $\lambda$-expansion $X^\lambda$ |
|---|---|
| $c_j$ | $c_j^\lambda = c_j/\lambda^{\omega_j - 1}$ |
| $\vec{y}(t)$ | $\vec{y}^\lambda(t)$ |
| $\vec{x}(k)$ | $\vec{x}^\lambda(k)$ |
| $\vec{z}(k)$ | $\vec{z}^\lambda(k)$ |
| $\vec{h}_j(\cdot)$ | $\vec{h}_j^\lambda(\cdot) = \vec{h}_j(\cdot)/\lambda^{\omega_j - 1}$ |
| $\vec{\mathfrak{p}}$ | $\vec{\mathfrak{p}}^\lambda = \lim_{k\to+\infty} \vec{p}^\lambda(\vec{x}^\lambda(k))$ |

In particular, the $\omega$-order reaction stochastic constants$\omega$ are divided by $\lambda^{\omega-1}$. The processes $\vec{y}^\lambda, \vec{x}^\lambda, \vec{z}^\lambda$ themselves correspond to dynamics obtained by multiplying all the initial quantities by $\lambda$ and by using the stochastic constants $c_j^\lambda$. Finally, studying the thermodynamics limits is equivalent with studying the case when $\lambda \to +\infty$.

**Proposition 3** (Expectancies at thermodynamics limit)**.** *Let $(\vec{x}(k))_{k\in\mathbb{N}}$ be the discretized dynamics of a system, $(\vec{x}^\lambda(k))_{k\in\mathbb{N}}$ the dynamics of its $\lambda$-expansion, and $\vec{z}^\lambda(k)$ the Bernoulli dynamics associated with $(\vec{x}^\lambda(k))_{k\in\mathbb{N}}$. Then the following holds:*

$$\forall k \in \mathbb{N},\ \|\mathbb{E}\vec{x}^\lambda(k) - \mathbb{E}\vec{z}^\lambda(k)\| \leq k\|S\| d_e^\lambda(\lambda\vec{x}_0) + o_{\lambda\to+\infty}(1).$$

*Proof.* From the properties of $\|\cdot\|$, we have:

$$\begin{aligned}\|\mathbb{E}\vec{x}^\lambda(k) - \mathbb{E}\vec{z}^\lambda(k)\| &= \left\| S \sum_{l=1}^{k} \left( \mathbb{E}(\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{\mathfrak{p}}^\lambda \right) \right\| \\ &\leq \|S\| \sum_{l=1}^{k} \left( \mathbb{E}\left( \|\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{\mathfrak{p}}^\lambda\| \right) \right).\end{aligned}$$

In addition, we have:

$$\begin{aligned}\|\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{\mathfrak{p}}^\lambda\| &= \|\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{\mathfrak{p}}^\lambda\| \\ &\leq \|\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{p}^\lambda(\vec{x}^\lambda(0))\| + \|\vec{p}^\lambda(\vec{x}^\lambda(0)) - \vec{\mathfrak{p}}^\lambda\| \\ &\leq \max_{\vec{w}\in B_{l-1}} \|\vec{p}^\lambda(\lambda\vec{x}(0) + \vec{w}) - \vec{p}^\lambda(\lambda\vec{x}(0))\| + d_e^\lambda(\vec{x}^\lambda(0)),\end{aligned}$$

where $B_{l-1}$ is the finite sets of moves than can be performed with $l-1$ reactions.

In order to estimate the error term, we need to prove that

$$p_j(\lambda\vec{y} + \vec{w}) - p_j(\lambda\vec{y}) \to 0. \tag{12}$$

To that goal, remind that

$$h_j^\lambda(\lambda\vec{y} + \vec{w}) = \frac{c_j}{\lambda^{\omega_j - 1}} \prod_{i=1}^{n} \binom{\lambda x_i + w_i}{\alpha_{i,j}}.$$

From the study of the dominant coefficient of this polynomial in $\lambda$ we get

$$\begin{aligned}\forall j = 1, \ldots, m, \quad h_j^\lambda(\lambda\vec{y} + \vec{w}) &\underset{\lambda\to+\infty}{\sim} \lambda c_j \prod_{i=1}^{n} \frac{x_i^{\alpha_{i,j}}}{\alpha_{i,j}!}, \\ h_0^\lambda(\lambda\vec{y} + \vec{w}) &\underset{\lambda\to+\infty}{\sim} \lambda \left( \sum_{j=1}^{m} c_j \prod_{i=1}^{n} \frac{x_i^{\alpha_{i,j}}}{\alpha_{i,j}!} \right).\end{aligned}$$

We deduce that $\lim_{\lambda\to\infty} h_0^\lambda(\lambda\vec{y} + \vec{w}) = +\infty$ for all $\vec{y}$, $\vec{w}$. Therefore, $p_0^\lambda(\lambda\vec{y} + \vec{w}) = \exp(-\delta t h_0^\lambda(\lambda\vec{y} + \vec{w}))$ and $p_0^\lambda(\lambda\vec{y})$ tend to 0 assuming that $\vec{w} = \vec{0}$. This implies that $p_0(\lambda\vec{y} + \vec{w}) - p_1(\lambda\vec{y}) \to 0$.

Let us now prove this formula for $j = 1, \ldots, m$:

$$\frac{p_j^\lambda(\lambda\vec{y} + \vec{w})}{p_j^\lambda(\lambda\vec{y})} = \frac{(1 - p_0^\lambda(\lambda\vec{y} + \vec{w})) \times h_j^\lambda(\lambda\vec{y} + \vec{w}) \times h_0^\lambda(\lambda\vec{y})}{(1 - p_0^\lambda(\lambda\vec{y})) \times h_j^\lambda(\lambda\vec{y}) \times h_0^\lambda(\lambda\vec{y} + \vec{w})}$$

when $p_j(\lambda\vec{y}) \neq 0$ (otherwise, the variables can be permuted to prove similar equation). We have already proved that $(1 - p_0^\lambda(\lambda\vec{y} + \vec{w})) \to 1$ and $(1 - p_0^\lambda(\lambda\vec{y})) \to 1$. In

addition, we have proved that $h_j^\lambda(\lambda\vec{y}+\vec{w})/h_j^\lambda(\lambda\vec{y}) \to 1$ and that $h_0^\lambda(\lambda\vec{y})/h_0^\lambda(\lambda\vec{y}+\vec{w}) \to 1$. Hence $\frac{p_j^\lambda(\lambda\vec{y}+\vec{w})}{p_j^\lambda(\lambda\vec{y})} \to 1$.

Coming back to the beginning of the proof, we deduce that $\|\vec{p}^\lambda(\vec{x}^\lambda(l-1)) - \vec{\mathfrak{p}}^\lambda\| \leq o_{\lambda\to+\infty}(1) + d_e^\lambda(\vec{x}^\lambda(0))$. Hence: $\|\mathbb{E}\vec{x}^\lambda(k) - \mathbb{E}\vec{z}^\lambda(k)\| \leq o_{\lambda\to+\infty}(1) + k\|S\|d_e^\lambda(\vec{x}^\lambda(0))$. □

We apply a similar two procedure to approximate co-variance matrices at their their thermodynamics limit.

**Proposition 4.** *The co-variance of $\vec{x}_k$ can be decomposed into three terms: (1) the co-variance of the corresponding Bernoulli dynamics, (2) a term related to the maintenance of equilibria, and (3) a term involving the initial distance to equilibria which is convergent when $k \to +\infty$.*

$$\begin{aligned}\operatorname{Cov}\vec{x}(k) = \operatorname{Cov}\vec{z}(k) &+ \sum_{l=0}^{k-1}(\operatorname{Cov}(\vec{x}(l),\vec{p}(\vec{x}(l)))S^t + S\operatorname{Cov}(\vec{p}(\vec{x}(l)),\vec{x}(l))) \\ &+ \sum_{l=0}^{k-1} S\left[\operatorname{diag}(\mathbb{E}\vec{p}(\vec{x}(l)) - \vec{\mathfrak{p}}) - (\mathbb{E}\vec{p}(\vec{x}(l))\mathbb{E}\vec{p}(\vec{x}(l))^t - \vec{\mathfrak{p}}\vec{\mathfrak{p}}^t)\right]S^t\end{aligned}$$

*Proof.* For any pair of random vectors $(\vec{y},\vec{x})$, let $\operatorname{Cov}(\vec{y},\vec{x}) = \mathbb{E}(\vec{y}\vec{x}^t) - \mathbb{E}\vec{y}\mathbb{E}\vec{x}^t$ and $\operatorname{Cov}(\vec{y}) = \operatorname{Cov}(\vec{y},\vec{y})$. Remind that co-variance is invariant under translation: $\operatorname{Cov}(\vec{y}+\vec{a},\vec{x}+\vec{b}) = \operatorname{Cov}(\vec{y},\vec{x})$. Therefore, for the sake of clarity, we assume that $\vec{x}(0) = 0$. Since $\vec{x}(k+1) = \vec{x}(k) + S\vec{e}_{\mu_{k+1}}$ and $\mathbb{E}(\vec{e}_{\mu_{k+1}}|\vec{x}(k)) = \vec{p}(\vec{x}(k))$, we have

$$\begin{aligned}\mathbb{E}(\vec{x}(k+1)\vec{x}(k+1)^t|\vec{x}(k)) &= \vec{x}(k)\vec{x}(k)^t + \vec{x}(k)\mathbb{E}((S\vec{e}_{\mu_{k+1}})^t|\vec{x}(k)) \\ &\quad + \mathbb{E}(S\vec{e}_{\mu_{k+1}}|\vec{x}(k))\vec{x}(k)^t + \mathbb{E}(S\vec{e}_{\mu_{k+1}}(S\vec{e}_{\mu_{k+1}})^t|\vec{x}(k)) \\ &= \vec{x}(k)\vec{x}(k)^t + \vec{x}(k)\vec{p}(\vec{x}(k))^t S^t \\ &\quad + S\vec{p}(\vec{x}(k))\vec{x}(k)^t + S\operatorname{diag}\vec{p}(\vec{x}(k))S^t.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}(\vec{x}(k+1)\vec{x}(k+1)^t) &= \mathbb{E}\mathbb{E}(\vec{x}(k+1)\vec{x}(k+1)^t|\vec{x}(k)) \\ &= \mathbb{E}(\vec{x}(k)\vec{x}(k)^t) + \mathbb{E}(\vec{x}(k)\vec{p}(\vec{x}(k))^t)S^t \\ &\quad + S\mathbb{E}(\vec{p}(\vec{x}(k))\vec{x}(k)^t) + S\operatorname{diag}\mathbb{E}\vec{p}(\vec{x}(k))S^t.\end{aligned}$$

In addition, we have that

$$\mathbb{E}\vec{x}(k+1) = \mathbb{E}\mathbb{E}(\vec{x}(k) + S\vec{e}_{\mu_{k+1}}|\vec{x}(k)) = \mathbb{E}y(k) + S\mathbb{E}\vec{p}(\vec{x}(k)).$$

This implies that:

$$\begin{aligned}\mathbb{E}\vec{x}(k+1)\mathbb{E}\vec{x}(k+1)^t = \mathbb{E}\vec{x}(k)\mathbb{E}\vec{x}(k)^t + \mathbb{E}\vec{x}(k)\mathbb{E}\vec{p}(\vec{x}(k))^t S^t \\ + S\mathbb{E}\vec{p}(\vec{x}(k))\mathbb{E}y(k)^t + S\mathbb{E}\vec{p}(\vec{x}(k))\mathbb{E}\vec{p}(\vec{x}(k))^t S^t.\end{aligned}$$

Substracting the two equalities yields:

$$\begin{aligned}\operatorname{Cov}(\vec{x}(k+1)) &= \operatorname{Cov}(\vec{x}(k)) + \left[\mathbb{E}(\vec{x}(k)\vec{p}(\vec{x}(k))^t) - \mathbb{E}\vec{x}(k)\mathbb{E}\vec{p}(\vec{x}(k))^t\right] S^t \\ &+ S\left[\mathbb{E}(\vec{p}(\vec{x}(k))\vec{x}(k)^t) - \mathbb{E}\vec{p}(\vec{x}(k))\mathbb{E}y(k)^t\right] \\ &+ S\left[\operatorname{diag}\mathbb{E}\vec{p}(\vec{x}(k)) - \mathbb{E}\vec{p}(\vec{x}(k))\mathbb{E}\vec{p}(\vec{x}(k))^t\right] S, \\ &= \operatorname{Cov}(\vec{x}(k)) + \operatorname{Cov}(\vec{x}(k), \vec{p}(\vec{x}(k)))S^t \\ &+ S\operatorname{Cov}(\vec{p}(\vec{x}(k)), \vec{x}(k)) \\ &+ S\left[\operatorname{diag}\mathbb{E}\vec{p}(\vec{x}(k)) - \mathbb{E}\vec{p}(\vec{x}(k))\mathbb{E}\vec{p}(\vec{x}(k))^t\right] S^t.\end{aligned}$$

Summing up both equations, we deduce that

$$\begin{aligned}\operatorname{Cov}(\vec{x}(k)) = \sum_{l=0}^{k-1} \big(&\operatorname{Cov}(\vec{x}(l), \vec{p}(\vec{x}(l)))S^t \\ &+ S\operatorname{Cov}(\vec{p}(\vec{x}(l)), \vec{x}(l)) \\ &+ S\left[\operatorname{diag}\mathbb{E}\vec{p}(\vec{x}(l)) - \mathbb{E}\vec{p}(\vec{x}(l))\mathbb{E}\vec{p}(\vec{x}(l))^t\right] S^t\big).\end{aligned}$$

Since we know from proposition 1 that $\operatorname{Cov}\vec{z}(k) = kS(\operatorname{diag}\vec{\mathfrak{p}} - \vec{\mathfrak{p}}\vec{\mathfrak{p}}^t)S^t$, we get the stated result. □

**Proposition 5** (Covariance and variance at thermodynamics limit)**.** *Let $(\vec{x}(k))_{k\in\mathbb{N}}$ be the discretized dynamics of a system, $(\vec{x}^\lambda(k))_{k\in\mathbb{N}}$ the dynamics of its $\lambda$-expansion, and let $\vec{z}^\lambda(k)$ be the Bernoulli dynamics associated to $(\vec{x}^\lambda(k))_{k\in\mathbb{N}}$. Then*

$$\forall k \in \mathbb{N},\ \|\operatorname{Cov}\vec{x}^\lambda(k) - \operatorname{Cov}\vec{z}^\lambda(k)\| \leq k(n+1)\|S\|\|S^t\|d_e^\lambda(\lambda\vec{x}_0) + o_{\lambda\to+\infty}(1).$$

*Proof.* Similar arguments as those used in the proof of proposition 2 yield

$$\|\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{p}^\lambda(\vec{x}^\lambda(0))\| \leq \max_{\vec{w}\in B_l} \|\vec{p}^\lambda(\lambda\vec{x}(0) + \vec{w}) - \vec{p}^\lambda(\lambda\vec{x}(0))\| \to_{\lambda\to+\infty} 0$$

where $B_l$ stands for the set of possible moves when using $l$ reactions.

From proposition 4, we have that

$$\operatorname{Cov}\vec{x}^\lambda(k) = \operatorname{Cov}\vec{z}^\lambda(k) + \sum_{l=0}^{k-1}(\operatorname{Cov}(\vec{x}^\lambda(l), \vec{p}^\lambda(\vec{x}^\lambda(l)))S^t + S\operatorname{Cov}(\vec{p}^\lambda(\vec{x}^\lambda(l)), \vec{x}^\lambda(l)))$$

$$+ \sum_{l=0}^{k-1} S\left[\operatorname{diag}(\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{\mathfrak{p}}^\lambda) - (\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))^t - \vec{\mathfrak{p}}^\lambda\vec{\mathfrak{p}}^{\lambda t})\right] S^t. \quad (13)$$

According to the invariance of co-variance by translation, we deduce that:

$$\mathrm{Cov}(\vec{x}^\lambda(l), \vec{p}^\lambda(\vec{x}^\lambda(l)) = \mathrm{Cov}(\vec{x}^\lambda(l), \vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{p}^\lambda(\vec{x}^\lambda(0))) \underset{\lambda \to +\infty}{\to} 0.$$

This allows us to prove that the first summation term in (13) is equal to $o(1)$ when $\lambda \to +\infty$.

For the second term in (13), we notice that

$$\begin{aligned}\|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{\mathfrak{p}}^\lambda\| &\leq \|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{p}^\lambda(\vec{x}^\lambda(0))\| + \|\vec{p}^\lambda(\vec{x}^\lambda(0)) - \vec{\mathfrak{p}}^\lambda\| \\ &\leq o_{\lambda\to+\infty}(1) + d_e^\lambda(\lambda\vec{x}(0)).\end{aligned}$$

We deduce that

$$\|\operatorname{diag}(\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{\mathfrak{p}}^\lambda)\| = \|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{\mathfrak{p}}^\lambda\| \leq o_{\lambda\to+\infty}(1) + d_e^\lambda(\lambda\vec{x}(0))$$

and that

$$\begin{aligned}&\|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))^t - \vec{\mathfrak{p}}^\lambda\vec{\mathfrak{p}}^{\lambda^t}\| \\ &\leq \|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l)) - \vec{\mathfrak{p}}^\lambda\| \underbrace{\|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))^t\|}_{\leq n} + \underbrace{\|\vec{\mathfrak{p}}^\lambda\|}_{\leq 1} \|\mathbb{E}\vec{p}^\lambda(\vec{x}^\lambda(l))^t - \vec{\mathfrak{p}}^{\lambda^t}\| \\ &\leq o_{\lambda\to+\infty}(1) + (n+1)d_e^\lambda(\lambda\vec{x}(0)).\end{aligned}$$

Triangular inequalities allow us to end up the proof.

□

From these results we deduce that the error terms for expectancy, variance and co-variance of a stochastic process and its Bernoulli approximation are neglectable si the system are close from the thermodynamics limit. Our approximation of expectancy, variance and co-variance of a stochastic process hold when the networks do have few enough compounds to make effective the stochastic effects of the system, and large enough compounds to enable the Bernoulli approximation.

## 0.2 A StochPy model

In this appendix, we provide both the model and the simulation script used to study the four models of section 4. Here, only the first model is given but a full set of models and simulation scripts are available in the companion website `https://github.com/jeremiebourdon/Simulations_IJCAI`.

### 0.2.1 Model 1 in a StochPy format

```
# Reactions
R1:
    A > B
    k1*A

R2:
    B + cof > C + byp + byp2
    k2*B

R3:
    B + cof > D + byp + byp2
    k3*B

R4:
    D > E + cof
    k4*D

R5:
    C + cof > D
    k5*C

R6:
    C > E
    k6*C

R7:
    E > A
    k7*E

#InitPar
k1 = 1.0
k2 = 1.0
k3 = 1.0
k4 = 1.0
k5 = 1.0
k6 = 1.0
k7 = 1.0

#InitVar
```

```
A = 10000
B = 10000
C = 10000
D = 10000
E = 10000
cof = 10000
byp = 0
byp2 = 0
```

### 0.2.2 Python script

```
import stochpy
import numpy
import matplotlib.pyplot as plt

# determining the steady state values for A, B, C, D and E components
file = 'Modele1.psc'

smod = stochpy.SSA()
smod.model_dir = '../Models'
smod.Model(file)
smod.ChangeInitialSpeciesCopyNumber("A",5000)
smod.ChangeInitialSpeciesCopyNumber("B",5000)
smod.ChangeInitialSpeciesCopyNumber("C",5000)
smod.ChangeInitialSpeciesCopyNumber("D",5000)
smod.ChangeInitialSpeciesCopyNumber("E",5000)

smod.DoStochSim(mode='time', end=20,trajectories=1)

data=smod.data_stochsim.species
c1=data[:,smod.data_stochsim.species_labels.index('A')]
A0=numpy.mean(c1[int(3*len(c1)/4):len(c1)])
c1=data[:,smod.data_stochsim.species_labels.index('B')]
B0=numpy.mean(c1[int(3*len(c1)/4):len(c1)])
c1=data[:,smod.data_stochsim.species_labels.index('C')]
C0=numpy.mean(c1[int(3*len(c1)/4):len(c1)])
c1=data[:,smod.data_stochsim.species_labels.index('D')]
D0=numpy.mean(c1[int(3*len(c1)/4):len(c1)])
c1=data[:,smod.data_stochsim.species_labels.index('E')]
E0=numpy.mean(c1[int(3*len(c1)/4):len(c1)])
```

```
nb_trajectories=1000
gridsize=500

smod = stochpy.SSA()
smod.Model(file)

# starting from a steady state
smod.ChangeInitialSpeciesCopyNumber("A",A0)
smod.ChangeInitialSpeciesCopyNumber("B",B0)
smod.ChangeInitialSpeciesCopyNumber("C",C0)
smod.ChangeInitialSpeciesCopyNumber("D",D0)
smod.ChangeInitialSpeciesCopyNumber("E",E0)
smod.DoStochSim(end=gridsize,trajectories=20)
smod.PlotSpeciesTimeSeries(species2plot=['byp','byp2'])
smod.GetRegularGrid(n_samples=50)
smod.PlotAverageSpeciesTimeSeries(species2plot=['byp','byp2'])
smod.PlotAverageSpeciesTimeSeries(species2plot=['A','B','C','D','E'])

# co-variance between byp and byp2
byp=smod.data_stochsim.species_labels.index('byp')
byp2=smod.data_stochsim.species_labels.index('byp2')

expesc1_100=numpy.empty([gridsize,nb_trajectories])
expesc2_100=numpy.empty([gridsize,nb_trajectories])

print ("Experiments for model 1")
for e in range (0,nb_trajectories):
    # starting from a steady state
    smod.ChangeInitialSpeciesCopyNumber("A",A0)
    smod.ChangeInitialSpeciesCopyNumber("B",B0)
    smod.ChangeInitialSpeciesCopyNumber("C",C0)
    smod.ChangeInitialSpeciesCopyNumber("D",D0)
    smod.ChangeInitialSpeciesCopyNumber("E",E0)
    smod.DoStochSim(end=gridsize,trajectories=1)
    data=smod.data_stochsim.species
    c1=data[:,byp]
    c2=data[:,byp2]
    for i in range (0,gridsize):
        expesc1_100[i,e]=c1[i]
        expesc2_100[i,e]=c2[i]
v1=[numpy.cov(expesc1_100[x,:],expesc2_100[x,:])[0,1]
```

```
        for x in range(0,gridsize)]

plt.show()
plt.title('Correlation between byp and byp2')
plt.xlabel('Time')
plt.ylabel('Covariance')
plt.plot(numpy.linspace(0, gridsize, gridsize),v1)
plt.show()
```