

Comparison Theorems for Stochastic Chemical Reaction Networks

Felipe A. Campos^{1,*}, Simone Bruno^{2,*}, Yi Fu¹, Domitilla Del Vecchio², and Ruth J. Williams¹

¹*Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla CA 92093-0112. Email: (fcamposv,yif064,rjwilliams)@ucsd.edu*

²*Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. Emails: (sbruno,ddv)@mit.edu*

**These authors contributed equally: F. A. Campos and S. Bruno*

Supplementary Information (SI) file

S.1 Criteria for Positive Recurrence and Exponential Ergodicity with Application to Examples 4.2 and 4.5

S.1.1 Foster-Lyapunov Conditions for Positive Recurrence and Exponential Ergodicity

Here we recall fairly general conditions for positive recurrence and exponential ergodicity of a continuous-time Markov chain. Such conditions are well known and are usually referred to as Foster-Lyapunov-type conditions. We also apply these to the Markov chains in Examples 4.2 and 4.5, in subsections S.1.2 and S.1.3.

Theorem S.1. *Let X be an irreducible continuous-time Markov chain with state space $\mathcal{X}$ and infinitesimal generator Q . Suppose $V : \mathcal{X} \rightarrow \mathbb{R}_+$ is norm-like, that is $\{x \in \mathcal{X} : V(x) \leq a\}$ is compact⁶ for each $a \in \mathbb{R}_+$. Further assume that for some $c > 0$, $d > 0$ and a compact set C ,*

$$QV(x) \leq -c + d\mathbf{1}_C(x), \quad \text{for all } x \in \mathcal{X}. \quad (\text{S.1})$$

Then, X is non-explosive and positive recurrent, and has a unique stationary distribution π . If instead of (S.1), we have that for some $c' > 0$ and $d' > 0$,

$$QV(x) \leq -c'V(x) + d', \quad \text{for all } x \in \mathcal{X}, \quad (\text{S.2})$$

then (S.1) automatically holds and the consequences stated above hold, and in addition, the stationary distribution satisfies

$$\pi(V) = \sum_{x \in \mathcal{X}} \pi_x V(x) < \infty,$$

and there is $0 < B < \infty$ and $0 < \beta < 1$ such that for all $t \geq 0$ and $x \in \mathcal{X}$,

$$\sum_{y \in \mathcal{X}} |P_{xy}(t) - \pi_y| \leq \|P_{x\bullet}(t) - \pi\|_{V+1} \leq B(V(x) + 1)\beta^t, \quad (\text{S.3})$$

⁶Since X is a Markov chain, $\mathcal{X}$ is finite or countable, and we endow it with the usual discrete topology consisting of all subsets of $\mathcal{X}$.

where $P_{xy}(t) = \mathbb{P}[X(t) = y | X(0) = x]$ and

$$\|P_{x\bullet}(t) - \pi\|_{V+1} = \sup_{|g| \leq V+1} \left| \sum_{y \in \mathcal{X}} (P_{xy}(t) - \pi_y) g(y) \right|.$$

Remark S.1. When the second inequality in (S.3) holds, we say that the Markov chain is exponentially ergodic in the $(V + 1)$ -norm.

Proof. We will verify the sufficient conditions for each of non-explosion, positive recurrence and exponential ergodicity given in Meyn & Tweedie [1]. Note that X is a Borel right-process under the definition in Sharpe [2]. In addition, since the state space is discrete, each compact set is finite and therefore petite.

For $m \in \mathbb{Z}_+$, if Q_m is the infinitesimal generator for the Markov chain X killed upon exit from $O_m = \{x \in \mathcal{X} : V(x) \leq m\}$ ⁷, then $Q_m V(x) \leq QV(x)$ for $x \in O_m$. It then follows from (S.1) that conditions (CD0) and (CD2) (with $f = 1$) in [1] hold with Q_m in place of $\mathcal{A}_m$ there. By Theorem 2.1 and Theorem 4.2 in [1], the Markov chain is non-explosive and positive recurrent, and it has a unique stationary distribution π .

On the other hand, if (S.2) holds, then (S.1) holds, using the norm like property of V . Furthermore, (S.2) implies that conditions (CD0), (CD2) (with $f = V + 1$) and (CD3) in [1] hold with Q_m in place of $\mathcal{A}_m$ there. By Theorem 2.1, Theorem 4.2 and Theorem 6.1 in [1], the Markov chain is non-explosive and positive recurrent, with a unique stationary distribution π such that $\pi(V) < \infty$, and it is exponentially ergodic in the $(V + 1)$ -norm, that is the second inequality in (S.3) holds for all $t \geq 0$ and $x \in \mathcal{X}$. For fixed $t \geq 0$ and $x \in \mathcal{X}$, setting $g(y) = \text{sgn}(P_{xy}(t) - \pi_y)$, for $y \in \mathcal{X}$, we have that $|g| \leq 1 \leq V + 1$, and

$$\sum_{y \in \mathcal{X}} |P_{xy}(t) - \pi_y| = \left| \sum_{y \in \mathcal{X}} (P_{xy}(t) - \pi_y) g(y) \right| \leq \|P_{x\bullet}(t) - \pi\|_{V+1},$$

yielding the first inequality in (S.3). $\square$

S.1.2 Application to Example 4.2

For Example 4.2, we first show that the Markov chain is irreducible. For this, consider $x^\circ = (0, 0, E_{\text{tot}}, 0)$ and any fixed state $x = (x_1, x_2, x_3, E_{\text{tot}} - x_3) : 0 \leq x_3 \leq E_{\text{tot}}$. Starting at x° , by having reaction ⑤ fire $x_1 + E_{\text{tot}} - x_3 + x_2$ times in succession, then having reaction ①, immediately followed by reaction ③, fire x_2 times in succession and then reaction ① fire $E_{\text{tot}} - x_3$ times, without any other reactions firing, we see that the Markov chain can transition with positive probability from x° to x . Since each reaction is reversible, it also follows that the Markov chain can transition from x to x° with positive probability. Thus, the Markov chain is irreducible.

Next we will introduce a norm-like function V and show that (S.1) holds. For each $x \in \mathcal{X}$, let

$$V(x) = x_1^2 + ((2E_{\text{tot}} - 1)b + 1)x_2 + bx_4^2,$$

⁷Upon exit from O_m , the killed process goes to a cemetery state Δ_m in $\mathcal{X} \setminus O_m$ where $V(\Delta_m) = \min\{V(x) : x \in \mathcal{X} \setminus O_m\}$.

where

$$b = \frac{1 + (\kappa_5 + \kappa_2 E_{\text{tot}} + \kappa_3 E_{\text{tot}}) + \frac{(2\kappa_5 + \kappa_6 + 2\kappa_2 E_{\text{tot}})^2}{8\kappa_6}}{\kappa_2 E_{\text{tot}}(2E_{\text{tot}} - 1)}. \quad (\text{S.4})$$

Notice that $b > 0$ since $E_{\text{tot}} \geq 1$. Then, for each $a \in \mathbb{R}_+$, $\{x \in \mathcal{X} : V(x) \leq a\}$ consists of finitely many states, and for any $x \in \mathcal{X}$,

$$\begin{aligned} QV(x) &= \sum_{j=1}^6 \Upsilon_j(x) \cdot (V(x + v_j) - V(x)) \\ &= \kappa_1 x_1 x_3 \cdot (((x_1 - 1)^2 + b(x_4 + 1)^2) - (x_1^2 + b x_4^2)) \\ &\quad + \kappa_2 x_4 \cdot (((x_1 + 1)^2 + b(x_4 - 1)^2) - (x_1^2 + b x_4^2)) \\ &\quad + \kappa_3 x_4 \cdot (((2E_{\text{tot}} - 1)b + 1)(x_2 + 1) + b(x_4 - 1)^2) - (((2E_{\text{tot}} - 1)b + 1)x_2 + b x_4^2)) \\ &\quad + \kappa_4 x_2 x_3 \cdot (((2E_{\text{tot}} - 1)b + 1)(x_2 - 1) + b(x_4 + 1)^2) - (((2E_{\text{tot}} - 1)b + 1)x_2 + b x_4^2)) \\ &\quad + \kappa_5 \cdot ((x_1 + 1)^2 - x_1^2) + \kappa_6 x_1 \cdot ((x_1 - 1)^2 - x_1^2) \\ &= \kappa_1 x_1 x_3 \cdot (-2x_1 + 1 + b(2x_4 + 1)) + \kappa_2 x_4 \cdot (2x_1 + 1 + b(-2x_4 + 1)) \\ &\quad + \kappa_3 x_4 \cdot (((2E_{\text{tot}} - 1)b + 1) + b(-2x_4 + 1)) \\ &\quad + \kappa_4 x_2 x_3 \cdot (-((2E_{\text{tot}} - 1)b + 1) + b(2x_4 + 1)) \\ &\quad + \kappa_5 \cdot (2x_1 + 1) + \kappa_6 x_1 \cdot (-2x_1 + 1) \\ &= -(2\kappa_1 x_3 + 2\kappa_6) \cdot x_1^2 + (2\kappa_5 + \kappa_6 + \kappa_1(1 + b)x_3 + 2b\kappa_1 x_3 x_4 + 2\kappa_2 x_4) \cdot x_1 \\ &\quad + \kappa_4(-2bE_{\text{tot}} x_3 + 2bx_3 x_4 + (2b - 1)x_3) \cdot x_2 \\ &\quad + \kappa_5 + (\kappa_2(1 + b) + \kappa_3(2bE_{\text{tot}} + 1))x_4 - 2b(\kappa_2 + \kappa_3)x_4^2 \\ &= -(2\kappa_1 x_3 + 2\kappa_6) \cdot x_1^2 + (2\kappa_5 + \kappa_6 + \kappa_1(1 + b)x_3 + 2b\kappa_1 x_3 x_4 + 2\kappa_2 x_4) \cdot x_1 \\ &\quad + \kappa_4(-2bx_3^2 + (2b - 1)x_3) \cdot x_2 + \kappa_5 + (\kappa_2(1 + b) + \kappa_3(2bE_{\text{tot}} + 1))x_4 - 2b(\kappa_2 + \kappa_3)x_4^2. \end{aligned}$$

The last equality uses the fact that $x_3 = E_{\text{tot}} - x_4$. For the following, we note that $-2bx_3^2 + (2b - 1)x_3 = (-2bx_3 + (2b - 1))x_3 \leq -1$ when $x_3 \geq 1$.

We now consider two cases for $QV(x)$: when $x_3 = 0$ and $x_3 > 0$. For the first case, when $x_3 = 0$, we have $x_4 = E_{\text{tot}}$ and

$$QV(x) = -2\kappa_6 \cdot x_1^2 + (2\kappa_5 + \kappa_6 + 2\kappa_2 E_{\text{tot}}) \cdot x_1 + \kappa_5 + (\kappa_2(1 + b) + \kappa_3)E_{\text{tot}} - 2b\kappa_2 E_{\text{tot}}^2.$$

As a quadratic function, the last expression is bounded above by

$$b\kappa_2 E_{\text{tot}}(1 - 2E_{\text{tot}}) + (\kappa_5 + \kappa_2 E_{\text{tot}} + \kappa_3 E_{\text{tot}}) + \frac{(2\kappa_5 + \kappa_6 + 2\kappa_2 E_{\text{tot}})^2}{8\kappa_6} = -1,$$

since b is chosen as in (S.4). For the second case when $x_3 \in \{1, \dots, E_{\text{tot}}\}$, we have

$$\begin{aligned} QV(x) &\leq -2\kappa_6 \cdot x_1^2 + (2\kappa_5 + \kappa_6 + \kappa_1(1 + b)x_3 + 2b\kappa_1 x_3 x_4 + 2\kappa_2 x_4) \cdot x_1 \\ &\quad - \kappa_4 \cdot x_2 + \kappa_5 + (\kappa_2(1 + b) + \kappa_3(2bE_{\text{tot}} + 1))x_4 - 2b(\kappa_2 + \kappa_3)x_4^2 \\ &\leq -2\kappa_6 \cdot x_1^2 + \kappa_6 \cdot x_1^2 + \left(\frac{2\kappa_5 + \kappa_6 + \kappa_1(1 + b)x_3 + 2b\kappa_1 x_3 x_4 + 2\kappa_2 x_4}{2\sqrt{\kappa_6}} \right)^2 \\ &\quad - \kappa_4 \cdot x_2 + \kappa_5 + (\kappa_2(1 + b) + \kappa_3(2bE_{\text{tot}} + 1))x_4 - 2b(\kappa_2 + \kappa_3)x_4^2 \\ &\leq -\kappa_6 \cdot x_1^2 - \kappa_4 \cdot x_2 + \frac{(2\kappa_5 + \kappa_6 + \kappa_1(1 + b)E_{\text{tot}} + 2b\kappa_1 E_{\text{tot}}^2 + 2\kappa_2 E_{\text{tot}})^2}{4\kappa_6} \\ &\quad + \kappa_5 + (\kappa_2(1 + b) + \kappa_3(2bE_{\text{tot}} + 1))E_{\text{tot}}, \end{aligned}$$

where we have used the fact that $2ab \leq a^2 + b^2$ for the second inequality and the last expression will be less than or equal to -1 whenever x_1 or x_2 is sufficiently large.

Let $C = \{x \in \mathcal{X} : QV(x) > -1\}$. Then, by the above, C consists of finitely many points, which implies that C is a compact set. Then, (S.1) holds with $c = 1$ and $d = 1 + \max_{x \in C} QV(x) \vee 0$. It follows by Theorem S.1 that the Markov chain in Example 4.2 is non-explosive and positive recurrent, and has a unique stationary distribution π .

S.1.3 Application to Example 4.5

For Example 4.5, we first check that the Markov chain is irreducible. For this, consider $x^\circ = (0, D_{\text{tot}}, 0)$ and any fixed state $x = (x_1, x_2, x_3) : 0 \leq x_1 + x_2 \leq D_{\text{tot}}$. Starting at x° , by having reaction ⑤ fire x_3 times, then reaction ③ fire $D_{\text{tot}} - x_2$ times, and finally, reaction ② fire x_1 times, without any other reactions firing, we see that the Markov chain can transition from x° to x with positive probability. For the reverse transition, by having reaction ⑥ fire x_3 times, then reaction ④ fire x_1 times, and finally, reaction ① fire $D_{\text{tot}} - x_2$ times, we see that the Markov chain can transition from x to x° with positive probability. Thus, the Markov chain is irreducible.

Next we will introduce a norm-like function V and show that (S.2) holds. For each $x \in \mathcal{X}$, let

$$V(x) = x_3.$$

Then $\{x \in \mathcal{X} : V(x) \leq a\} = \{x \in \mathbb{Z}_+^3 : x_1 + x_2 \leq D_{\text{tot}}, x_3 \leq a\}$ consists of finitely many states for each $a \in \mathbb{R}_+$, and for any $x \in \mathcal{X}$,

$$\begin{aligned} QV(x) &= \sum_{j=1}^6 \Upsilon_j(x) \cdot (V(x + v_j) - V(x)) = \sum_{j=1}^6 \Upsilon_j(x) \cdot V(v_j) \\ &= \Upsilon_1(x) \cdot 0 + \Upsilon_2(x) \cdot 0 + \Upsilon_3(x) \cdot 0 + \Upsilon_4(x) \cdot 0 + \kappa_{5a}x_2 \cdot 1 + \kappa_{6a}x_3 \cdot (-1) \\ &\leq -\kappa_{6a}x_3 + \kappa_{5a}D_{\text{tot}} = -c'V(x) + d', \end{aligned}$$

where $c' = \kappa_{6a}$ and $d' = \kappa_{5a}$. Therefore, we conclude by Theorem S.1 that the Markov chain in Example 4.5 is non-explosive and positive recurrent with a unique stationary distribution π such that $\pi(V) < \infty$, and it is exponentially ergodic in the $(V + 1)$ -norm.

S.2 Derivation of Markov chain transition directions, v_j

S.2.1 Example 4.2

The set of reactions associated to the chemical reaction system in Fig. 3(a) is given by $\mathcal{R} = \{(v_1^-, v_1^+), (v_2^-, v_2^+), (v_3^-, v_3^+), (v_4^-, v_4^+), (v_5^-, v_5^+), (v_6^-, v_6^+)\}$, where (v_1^-, v_1^+) , (v_2^-, v_2^+) , (v_3^-, v_3^+) are defined as in Example 4.1, and $v_4^- = (0, 1, 1, 0)^T$, $v_4^+ = (0, 0, 0, 1)^T$, $v_5^- = v_6^+ = (0, 0, 0, 0)^T$, $v_5^+ = v_6^- = (1, 0, 0, 0)^T$. Then, the potential transitions of the Markov chain are in six possible directions, $v_j = v_j^+ - v_j^-$ for $j = 1, \dots, 6$, where $v_1 = -v_2 = (-1, 0, -1, 1)^T$, $v_3 = -v_4 = (0, 1, 1, -1)^T$, and $v_5 = -v_6 = (1, 0, 0, 0)^T$.

S.2.2 Example 4.3

The set of five reactions associated to the chemical reaction system in Fig. 4(a) is given by $\mathcal{R} = \{(v_1^-, v_1^+), (v_2^-, v_2^+), (v_3^-, v_3^+), (v_4^-, v_4^+), (v_5^-, v_5^+)\}$, where $v_1^- = v_3^- = (1, 0, 0, 0)^T$, $v_1^+ = v_2^- = v_5^- = (0, 1, 0, 0)^T$, $v_3^+ = v_4^- = v_5^+ = (0, 0, 1, 0)^T$, $v_2^+ = v_4^+ = (0, 0, 0, 1)^T$. Then, the potential transitions of the Markov chain are in five possible directions, $v_j = v_j^+ - v_j^-$, $j =$

1, ..., 5, where $v_1 = (-1, 1, 0, 0)^T$, $v_2 = (0, -1, 0, 1)^T$, $v_3 = (-1, 0, 1, 0)^T$, $v_4 = (0, 0, -1, 1)^T$ and $v_5 = (0, -1, 1, 0)^T$.

S.3 A generalization of Theorem 3.3

Here, we provide a more general version of Theorem 3.3. The simpler form given as Theorem 3.3 in the main text was used there because it is more straightforward to state and the conditions are easier to verify. However, the more general version of the theorem provided in this section can be useful in some cases, such as Example 4.3 (see Section S.3.2).

The generalization relies on the idea that grouping of vectors can be more general than what is described in (3.12), and so we introduce the following assumption.

Assumption S.1. Consider a collection of distinct vectors $v_1, \dots, v_n$ in $\mathbb{Z}^d \setminus \{0\}$ and a matrix $A \in \mathbb{Z}^{m \times d}$ with non-zero rows. Suppose that there exists a partition⁸ $\{G^1, \dots, G^s\}$ of $\{1, \dots, n\}$ and an associated bijection $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ such that for $p_0 := 0$ and $1 \leq k \leq s$, with $p_k := \sum_{\ell=1}^k |G^\ell|$, we have $\sigma(q) \in G^k$ for $p_{k-1} + 1 \leq q \leq p_k$, and whenever $p_{k-1} + 2 \leq q \leq p_k$, we also have for each $1 \leq i \leq m$ that either $\langle A_{i\bullet}, v_{\sigma(q)} \rangle$ is equal to $\langle A_{i\bullet}, v_{\sigma(q-1)} \rangle$ or it is 0.

Theorem S.2. Consider a non-empty set $\mathcal{X} \subseteq \mathbb{Z}_+^d$, suppose that Assumption S.1 holds and consider two collections of non-negative functions on $\mathcal{X}$, $\Upsilon = (\Upsilon_1, \dots, \Upsilon_n)$ and $\check{\Upsilon} = (\check{\Upsilon}_1, \dots, \check{\Upsilon}_n)$, such that (3.3) holds and the associated continuous-time Markov chains do not explode in finite time. Further suppose that both of the following conditions hold:

- (i) For each $1 \leq j \leq n$, the vector Av_j has entries in $\{-1, 0, 1\}$ only.
- (ii) For each $x \in \mathcal{X}$, $1 \leq i \leq m$ and $y \in \partial_i(K_A + x) \cap \mathcal{X}$ we have that for each $1 \leq k \leq s$,

$$\sum_{j \in G_i^{k,-}} \check{\Upsilon}_j(y) \leq \sum_{j \in G_i^{k,-}} \Upsilon_j(x), \quad \text{where } G_i^{k,-} = \{j \in G^k \mid \langle A_{i\bullet}, v_j \rangle = -1\},$$

and

$$\sum_{j \in G_i^{k,+}} \check{\Upsilon}_j(y) \geq \sum_{j \in G_i^{k,+}} \Upsilon_j(x), \quad \text{where } G_i^{k,+} = \{j \in G^k \mid \langle A_{i\bullet}, v_j \rangle = 1\}.$$

Then, for each pair $x^\circ, \check{x}^\circ \in \mathcal{X}$ such that $x^\circ \preceq_A \check{x}^\circ$, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with two continuous-time Markov chains $X = \{X(t), t \geq 0\}$ and $\check{X} = \{\check{X}(t), t \geq 0\}$ defined there, each having state space $\mathcal{X} \subseteq \mathbb{Z}_+^d$, with infinitesimal generators Q and $\check{Q}$, associated with Υ and $\check{\Upsilon}$ respectively, with initial conditions $X(0) = x^\circ$ and $\check{X}(0) = \check{x}^\circ$ and such that:

$$\mathbb{P} \left[X(t) \preceq_A \check{X}(t) \text{ for every } t \geq 0 \right] = 1. \quad (\text{S.5})$$

Remark S.2. Under Assumption S.1, for a given $1 \leq k \leq s$ and $1 \leq i \leq m$, at most one of $G_i^{k,-}$ and $G_i^{k,+}$ is non-empty. If $G_i^{k,-} \neq \emptyset$, then $G_i^{k,-} = \{\sigma(q) \mid p_{k-1} + 1 \leq q \leq q^*\}$ where $q^* = \max\{q \mid p_{k-1} + 1 \leq q \leq p_k \text{ and } \langle A_{i\bullet}, v_{\sigma(q)} \rangle = -1\}$. On the other hand, if $G_i^{k,+} \neq \emptyset$, then $G_i^{k,+} = \{\sigma(q) \mid p_{k-1} + 1 \leq q \leq q^*\}$ where $q^* = \max\{q \mid p_{k-1} + 1 \leq q \leq p_k \text{ and } \langle A_{i\bullet}, v_{\sigma(q)} \rangle = 1\}$. Furthermore, in either case, for $q^* < q \leq p_k$, we have $\langle A_{i\bullet}, v_{\sigma(q)} \rangle = 0$.

⁸In particular, $\{G^1, \dots, G^s\}$ is a finite collection of non-empty disjoint sets of distinct numbers, the union of which is $\{1, \dots, n\}$.

The proof for Theorem S.2 can be found in Section S.3.1. With Theorem S.2 in place, we can extend Theorems 3.4 and 3.5 by adding an additional alternative condition (iv): Assumption S.1 holds, and conditions (i) and (ii) in Theorem S.2 are satisfied. If this is satisfied instead of one of (i) – (iii) in Theorems 3.4 and 3.5, then the conclusions of these theorems about (mean) first passage times and stationary distributions will still hold.

S.3.1 Proof of Theorem S.2

We initially assume that $\sup_{x \in \mathcal{X}} \Upsilon_j(x) < \infty$ and $\sup_{x \in \mathcal{X}} \check{\Upsilon}_j(x) < \infty$ for every $1 \leq j \leq n$, and let $\lambda > 0$ such that (5.2) holds. We shall relax these assumptions later. Further suppose that Assumption S.1 and condition (i) of Theorem S.2 both hold. For $x \in \mathcal{X}$, define $I^k(x)$, $I_q^k(x)$, Ψ_λ , $\check{I}_q^k(x)$, $\check{I}_q^k(x)$ and $\check{\Psi}_\lambda$ in the same manner as in the proof of Theorem 3.3 (see (5.32) – (5.34)), with $\{G^k \mid 1 \leq k \leq s\}$, $\{p_k \mid 0 \leq k \leq s\}$ and σ as in Assumption S.1. Our proof of Theorem S.2 has some elements that are the same as those for the proof of Theorem 3.3. However, some additional elements are needed. We give the details for completeness. As for Theorem 3.3, $\Psi_\lambda(\cdot, \cdot)$ and $\check{\Psi}_\lambda(\cdot, \cdot)$ are well-defined as $\mathcal{X}$ -valued functions.

Lemma S.1. *Suppose that $x, y \in \mathcal{X}$ are such that $x \preceq_A y$ and the following hold: whenever $y \in \partial_i(K_A + x) \cap \mathcal{X}$ for some $1 \leq i \leq m$, we have that for each $1 \leq k \leq s$,*

$$\sum_{j \in G_i^{k,-}} \check{\Upsilon}_j(y) \leq \sum_{j \in G_i^{k,-}} \Upsilon_j(x), \quad \text{where } G_i^{k,-} = \{j \in G^k \mid \langle A_{i\bullet}, v_j \rangle = -1\}, \quad (\text{S.6})$$

and

$$\sum_{j \in G_i^{k,+}} \check{\Upsilon}_j(y) \geq \sum_{j \in G_i^{k,+}} \Upsilon_j(x), \quad \text{where } G_i^{k,+} = \{j \in G^k \mid \langle A_{i\bullet}, v_j \rangle = 1\}. \quad (\text{S.7})$$

Then, for each $u \in [0, 1]$:

$$\Psi_\lambda(x, u) \preceq_A \check{\Psi}_\lambda(y, u). \quad (\text{S.8})$$

Proof. First, we note that $\Psi_\lambda, \check{\Psi}_\lambda$ have the following properties: for every $u \in [0, 1]$, $1 \leq k \leq s$, $j \in G^k$,

$$\text{if } \Psi_\lambda(x, u) = x + v_j, \text{ then } \check{\Psi}_\lambda(y, u) \in \{y + v_\ell : \ell \in G^k\} \cup \{y\}, \quad (\text{S.9})$$

since $I_{\sigma^{-1}(j)}^k(x), \check{I}_{\sigma^{-1}(\ell)}^k(y) \subseteq [\frac{p_{k-1}}{n}, \frac{p_k}{n})$ for $\ell \in G^k$. Similarly,

$$\text{if } \check{\Psi}_\lambda(y, u) = y + v_j, \text{ then } \Psi_\lambda(x, u) \in \{x + v_\ell : \ell \in G^k\} \cup \{x\}. \quad (\text{S.10})$$

Furthermore, for $1 \leq k \leq s$, $1 \leq i \leq m$, $j \in G_i^{k,+}$, if $\sum_{\ell \in G_i^{k,+}} \check{\Upsilon}_\ell(y) \geq \sum_{\ell \in G_i^{k,+}} \Upsilon_\ell(x)$, then

$$\Psi_\lambda(x, u) = x + v_j \text{ implies that } \check{\Psi}_\lambda(y, u) = y + v_\ell \text{ for some } \ell \in G_i^{k,+}, \quad (\text{S.11})$$

since under the condition, we have $\cup_{q=p_{k-1}+1}^{q^*} I_q^k(x) \subseteq \cup_{q=p_{k-1}+1}^{q^*} \check{I}_q^k(y)$ where $q^* = \max\{q \mid p_{k-1} + 1 \leq q \leq p_k \text{ and } \langle A_{i\bullet}, v_{\sigma(q)} \rangle = 1\}$ and, by Assumption S.1, $G_i^{k,+} = \{\sigma(q) \mid p_{k-1} + 1 \leq q \leq q^*\}$. Similarly, for $1 \leq k \leq s$, $1 \leq i \leq m$, $j \in G_i^{k,-}$, if $\sum_{\ell \in G_i^{k,-}} \check{\Upsilon}_\ell(y) \leq \sum_{\ell \in G_i^{k,-}} \Upsilon_\ell(x)$, then

$$\check{\Psi}_\lambda(y, u) = y + v_j \text{ implies that } \Psi_\lambda(x, u) = x + v_\ell \text{ for some } \ell \in G_i^{k,-}. \quad (\text{S.12})$$

We also have that, for $1 \leq k \leq s$ and $j \in G^k$, $x \preceq_A y + v_j$ if and only if

$$\langle A_{i\bullet}, y - x \rangle + \langle A_{i\bullet}, v_j \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m. \quad (\text{S.13})$$

Similarly, $x + v_j \preceq_A y$ if and only if

$$\langle A_{i\bullet}, y - x \rangle - \langle A_{i\bullet}, v_j \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m. \quad (\text{S.14})$$

Furthermore, for $1 \leq k \leq s$ and $j, \ell \in G^k$, $x + v_\ell \preceq_A y + v_j$ if and only if

$$\langle A_{i\bullet}, y - x \rangle + \langle A_{i\bullet}, v_j - v_\ell \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m. \quad (\text{S.15})$$

To prove (S.8), we first consider the situation where $y \in \text{int}(K_A + x) = \{w \in \mathbb{R}^d \mid Ax < Aw\}$. Then, for each $1 \leq i \leq m$, $\langle A_{i\bullet}, y - x \rangle > 0$ and since $A \in \mathbb{Z}^{m \times d}$ and $y - x \in \mathbb{Z}^d$, we have $\langle A_{i\bullet}, y - x \rangle \geq 1$. This implies that for each $1 \leq k \leq s$ and $j \in G^k$,

$$\langle A_{i\bullet}, y - x \rangle + \langle A_{i\bullet}, v_j \rangle \geq 1 + \langle A_{i\bullet}, v_j \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m, \quad (\text{S.16})$$

since $\langle A_{i\bullet}, v_j \rangle \in \{-1, 0, 1\}$ by condition (i) of Theorem S.2. Similarly, for each $1 \leq k \leq s$ and $j \in G^k$,

$$\langle A_{i\bullet}, y - x \rangle - \langle A_{i\bullet}, v_j \rangle \geq 1 - \langle A_{i\bullet}, v_j \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m. \quad (\text{S.17})$$

In addition, for $1 \leq k \leq s$ and $j, \ell \in G^k$,

$$\langle A_{i\bullet}, y - x \rangle + \langle A_{i\bullet}, v_j - v_\ell \rangle \geq 1 + \langle A_{i\bullet}, v_j - v_\ell \rangle \geq 0, \quad \text{for every } 1 \leq i \leq m, \quad (\text{S.18})$$

since, by Assumption S.1, if $\langle A_{i\bullet}, v_j \rangle \neq 0$, then either $\langle A_{i\bullet}, v_\ell \rangle = \langle A_{i\bullet}, v_j \rangle$ or $\langle A_{i\bullet}, v_\ell \rangle = 0$. It follows from (S.16) – (S.18) that if $y \in \text{int}(K_A + x) \cap \mathcal{X}$, then for any $1 \leq k \leq s$ and $j, \ell \in G^k$:

$$x \preceq_A y + v_j, \quad x + v_j \preceq_A y \quad \text{and} \quad x + v_\ell \preceq_A y + v_j. \quad (\text{S.19})$$

We also have, by assumption, that $x \preceq_A y$. It follows that if $y \in \text{int}(K_A + x) \cap \mathcal{X}$, then $\{x, x + v_\ell \mid \ell \in G^k\} \preceq_A \{y, y + v_j \mid j \in G^k\}$ for $1 \leq k \leq s$ and consequently (S.8) holds for all $u \in [0, 1]$.

Now, we turn to the other situation where $y \in \partial_i(K_A + x) \cap \mathcal{X}$ for some $1 \leq i \leq m$. Then $\mathbf{K}_y := \{i \mid \langle A_{i\bullet}, y \rangle = \langle A_{i\bullet}, x \rangle, 1 \leq i \leq m\}$ is non-empty. Let $u \in [0, 1]$. We consider two cases.

Case 1: $\check{\Psi}_\lambda(y, u) = y + v_j$ for some $1 \leq j \leq n$.

Fix such an index j . Consider the unique $1 \leq k \leq s$ such that $j \in G^k$. Then, by (S.10), either $\Psi_\lambda(x, u) = x + v_\ell$ for some $\ell \in G^k$, or $\Psi_\lambda(x, u) = x$.

a) Suppose $\Psi_\lambda(x, u) = x + v_\ell$ for some $\ell \in G^k$. Observe that for every $i \notin \mathbf{K}_y$, $\langle A_{i\bullet}, y - x \rangle > 0$ and as for (S.18), $\langle A_{i\bullet}, (y + v_j) - (x + v_\ell) \rangle \geq 0$, while for $i \in \mathbf{K}_y$, $\langle A_{i\bullet}, (y + v_j) - (x + v_\ell) \rangle = \langle A_{i\bullet}, v_j \rangle - \langle A_{i\bullet}, v_\ell \rangle$. For each $i \in \mathbf{K}_y$,

i) if $\langle A_{i\bullet}, v_j \rangle = -1$ and $\langle A_{i\bullet}, v_\ell \rangle = 0$, then $j \in G_i^{k,-}$ and $\ell \notin G_i^{k,-}$. By (S.6), we would then have $\sum_{r \in G_i^{k,-}} \check{\Upsilon}_r(y) \leq \sum_{r \in G_i^{k,-}} \Upsilon_r(x)$, which would imply by (S.12) that $\Psi_\lambda(x, u) = x + v_r$ for some $r \in G_i^{k,-}$. Since we are assuming that $\Psi_\lambda(x, u) = x + v_\ell$ and we know the vectors $v_1, \dots, v_n$ are distinct, we obtain that $\ell = r \in G_i^{k,-}$. This contradicts $\ell \notin G_i^{k,-}$.

ii) if $\langle A_{i\bullet}, v_j \rangle = 0$ and $\langle A_{i\bullet}, v_\ell \rangle = 1$, then $j \notin G_i^{k,+}$ and $\ell \in G_i^{k,+}$. By (S.7), we would then have $\sum_{r \in G_i^{k,+}} \check{\Upsilon}_r(y) \geq \sum_{r \in G_i^{k,+}} \Upsilon_r(x)$, which would imply by (S.11) that $j \in G_i^{k,+}$. This contradicts $j \notin G_i^{k,+}$.

iii) in all the other cases, that is when

$$(\langle A_{i\bullet}, v_j \rangle, \langle A_{i\bullet}, v_\ell \rangle) \in \{(1, 0), (0, -1), (1, 1), (0, 0), (-1, -1)\},$$

$$\text{we have } \langle A_{i\bullet}, (y + v_j) - (x + v_\ell) \rangle \geq 0.$$

Combining the above for case a), we see that $\langle A_{i\bullet}, (y + v_j) - (x + v_\ell) \rangle \geq 0$ for each $1 \leq i \leq m$, which implies that $\Psi_\lambda(x, u) = x + v_\ell \preceq_A y + v_j = \check{\Psi}_\lambda(y, u)$.

b) Suppose $\Psi_\lambda(x, u) = x$. We claim that $y + v_j \in K_A + x$. To see this, observe that for every $i \notin \mathbf{K}_y$, $\langle A_{i\bullet}, y - x \rangle > 0$ and as for (S.16), $\langle A_{i\bullet}, (y + v_j) - x \rangle \geq 0$, while for $i \in \mathbf{K}_y$, $\langle A_{i\bullet}, (y + v_j) - x \rangle = \langle A_{i\bullet}, v_j \rangle \in \{-1, 0, 1\}$. For each $i \in \mathbf{K}_y$, if $\langle A_{i\bullet}, v_j \rangle = -1$, which means $j \in G_i^{k,-}$, then by (S.6) we would have $\sum_{\ell \in G_i^{k,-}} \check{\Upsilon}_\ell(y) \leq \sum_{\ell \in G_i^{k,-}} \Upsilon_\ell(x)$, which would imply by (S.12) that $\Psi_\lambda(x, u) = x + v_\ell$ for some $\ell \in G_i^{k,-}$, but this contradicts the assumption that $\Psi_\lambda(x, u) = x$. So we must have $\langle A_{i\bullet}, v_j \rangle \geq 0$ and hence $\langle A_{i\bullet}, (y + v_j) - x \rangle \geq 0$ for all $i \in \mathbf{K}_y$. Thus, $y + v_j \in K_A + x$ and so $\Psi_\lambda(x, u) = x \preceq_A y + v_j = \check{\Psi}_\lambda(y, u)$.

Case 2: $\check{\Psi}_\lambda(y, u) = y$. Again, we consider two subcases.

a) If $\Psi_\lambda(x, u) = x$, then (S.8) holds, because $x \preceq_A y$.

b) If $\Psi_\lambda(x, u) = x + v_j$ for some $1 \leq j \leq n$, we claim that $y \in K_A + x + v_j$ for the corresponding value of j . To see this, fix the value of j for which $\Psi_\lambda(x, u) = x + v_j$, let $1 \leq k \leq s$ be such that $j \in G^k$, and observe that for every $i \notin \mathbf{K}_y$, $\langle A_{i\bullet}, y - x \rangle > 0$ and as for (S.17), $\langle A_{i\bullet}, y - (x + v_j) \rangle \geq 0$, while for $i \in \mathbf{K}_y$, $\langle A_{i\bullet}, y - (x + v_j) \rangle = -\langle A_{i\bullet}, v_j \rangle \in \{-1, 0, 1\}$. For each $i \in \mathbf{K}_y$, if $\langle A_{i\bullet}, v_j \rangle = 1$, which means $j \in G_i^{k,+}$, then by (S.7), we would have $\sum_{\ell \in G_i^{k,+}} \check{\Upsilon}_\ell(y) \geq \sum_{\ell \in G_i^{k,+}} \Upsilon_\ell(x)$, which would imply by (S.11) that $\check{\Psi}_\lambda(y, u) = y + v_\ell$ for some $\ell \in G_i^{k,+}$, but this contradicts the assumption that $\check{\Psi}_\lambda(y, u) = y$. So we must have $\langle A_{i\bullet}, v_j \rangle \leq 0$ and hence $\langle A_{i\bullet}, y - (x + v_j) \rangle \geq 0$ for all $i \in \mathbf{K}_y$. Thus, we have $y \in K_A + x + v_j$ and then $\Psi_\lambda(x, u) = x + v_j \preceq_A y = \check{\Psi}_\lambda(y, u)$. $\square$

In order to prove Theorem S.2, from here on we can follow a similar procedure to the one used in the proof of Theorem 3.1 after Lemma 5.1 was proved there. For the case where (5.1) holds, we define two discrete-time processes, $Y = (Y_k)_{k \geq 0}$ and $\check{Y} = (\check{Y}_k)_{k \geq 0}$, by defining $Y_0 := x^\circ$, $\check{Y}_0 := \check{x}^\circ$, and for $k \geq 0$,

$$Y_{k+1} := \Psi_\lambda(Y_k, U_{k+1}), \quad \check{Y}_{k+1} := \check{\Psi}_\lambda(\check{Y}_k, U_{k+1}), \quad (\text{S.20})$$

and define X and $\check{X}$ using these and an independent Poisson process N as in (5.13). For the case where (5.1) does not hold, we can use a truncation procedure similar to that for Theorem 3.1. In both cases, we use Lemma S.1 instead of Lemma 5.1.

S.3.2 Two other A matrices for Example 4.3

Let

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & -1 & 0 \end{bmatrix}. \quad (\text{S.21})$$

For $x \in \mathcal{X}$, consider infinitesimal transition rates $\check{\Upsilon}_1(x), \check{\Upsilon}_2(x), \check{\Upsilon}_3(x), \check{\Upsilon}_4(x)$ and $\check{\Upsilon}_5(x)$ defined as for $\Upsilon_1(x), \Upsilon_2(x), \Upsilon_3(x), \Upsilon_4(x)$ and $\Upsilon_5(x)$ in (4.7), but with $\check{\kappa}_i$ in place of κ_i where $\check{\kappa}_i = \kappa_i$, for $i = 1, 2, 3, 4$, and $\check{\kappa}_5 \leq \kappa_5$. Suppose that $\kappa_2 > \kappa_4$. Now, let us verify that the assumptions of Theorem S.2 hold. Condition (i) holds since $Av_1 = (1, 0, -1)^T$, $Av_2 = (0, 0, 1)^T$, $Av_3 = (1, -1, -1)^T$, $Av_4 = (0, 1, 1)^T$ and $Av_5 = (0, -1, 0)^T$. Assumption S.1 holds with $G^1 = \{3, 1\}$, $G^2 = \{4, 2\}$, $G^3 = \{5\}$ and $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 4$, $\sigma(4) = 2$, $\sigma(5) = 5$. To verify that condition (ii) of Theorem S.2 holds, fix $x \in \mathcal{X}$ and first consider $y \in \partial_1(K_A + x) \cap \mathcal{X}$, where $\partial_1(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 = w_1, x_3 \geq w_3, x_2 + x_3 \geq w_2 + w_3, x_4 \leq w_4\}$. Given that $\langle A_{1\bullet}, v_1 \rangle = \langle A_{1\bullet}, v_3 \rangle = 1$, we need to check that $\Upsilon_1(x) + \Upsilon_3(x) \leq \check{\Upsilon}_1(y) + \check{\Upsilon}_3(y)$. Since $y \in \partial_1(K_A + x) \cap \mathcal{X}$, then $\Upsilon_1(x) = \kappa_1 x_1 = \kappa_1 y_1 = \check{\kappa}_1 y_1 = \check{\Upsilon}_1(y)$ and $\Upsilon_3(x) = \kappa_3 x_1 = \kappa_3 y_1 = \check{\kappa}_3 y_1 = \check{\Upsilon}_3(y)$, and so the desired inequality holds with equality. Secondly, consider $y \in \partial_2(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 \geq w_1, x_3 = w_3, x_2 \geq w_2, x_4 \leq w_4\}$. Given that $\langle A_{2\bullet}, v_3 \rangle = \langle A_{2\bullet}, v_5 \rangle = -1$ and $\langle A_{2\bullet}, v_4 \rangle = 1$, we need to check that $\Upsilon_3(x) \geq \check{\Upsilon}_3(y)$, $\Upsilon_4(x) \leq \check{\Upsilon}_4(y)$ and $\Upsilon_5(x) \geq \check{\Upsilon}_5(y)$. Since $y \in \partial_2(K_A + x) \cap \mathcal{X}$, then $\Upsilon_3(x) = \kappa_3 x_1 \geq \kappa_3 y_1 = \check{\kappa}_3 y_1 = \check{\Upsilon}_3(y)$, $\Upsilon_4(x) = \kappa_4 x_3 = \kappa_4 y_3 = \check{\kappa}_4 y_3 = \check{\Upsilon}_4(y)$ and $\Upsilon_5(x) = \kappa_5 x_2 \geq \kappa_5 y_2 \geq \check{\kappa}_5 y_2 = \check{\Upsilon}_5(y)$, and so the desired inequality holds. Lastly, consider $y \in \partial_3(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 \geq w_1, x_3 \geq w_3, x_2 + x_3 = w_2 + w_3, x_4 \leq w_4\}$. Given that $\langle A_{1\bullet}, v_1 \rangle = \langle A_{1\bullet}, v_3 \rangle = -1$ and $\langle A_{1\bullet}, v_2 \rangle = \langle A_{1\bullet}, v_4 \rangle = 1$, we need to check that $\Upsilon_1(x) + \Upsilon_3(x) \geq \check{\Upsilon}_1(y) + \check{\Upsilon}_3(y)$ and $\Upsilon_2(x) + \Upsilon_4(x) \leq \check{\Upsilon}_2(y) + \check{\Upsilon}_4(y)$. For $y \in \partial_3(K_A + x) \cap \mathcal{X}$, since $\kappa_2 > \kappa_4$ was assumed, we have that $\Upsilon_2(x) + \Upsilon_4(x) = \kappa_2 x_2 + \kappa_4 x_3 = (\kappa_2 - \kappa_4)x_2 + \kappa_4(x_2 + x_3) \leq (\kappa_2 - \kappa_4)y_2 + \kappa_4(y_2 + y_3) = \check{\kappa}_2 y_2 + \check{\kappa}_4 y_3 = \check{\Upsilon}_2(y) + \check{\Upsilon}_4(y)$ and $\Upsilon_1(x) = \kappa_1 x_1 \geq \kappa_1 y_1 = \check{\kappa}_1 y_1 = \check{\Upsilon}_1(y)$, $\Upsilon_3(x) = \kappa_3 x_1 \geq \kappa_3 y_1 = \check{\kappa}_3 y_1 = \check{\Upsilon}_3(y)$. Thus, the conditions of Theorem S.2 are satisfied and so the conclusion of that theorem holds.

Let $\Gamma = \{(0, 0, 0, S_{\text{tot}})\}$. This is an increasing set in $\mathcal{X}$ with respect to the relation $\preceq_A$. Let $T_{(0,0,0,S_{\text{tot}})}$, respectively $\check{T}_{(0,0,0,S_{\text{tot}})}$ be the first time that the Markov chain X , respectively $\check{X}$, reaches the set Γ . Then, by the generalization of Theorem 3.4, if $X(0) = \check{X}(0) = (S_{\text{tot}}, 0, 0, 0)$, we have that $\check{T}_{(0,0,0,S_{\text{tot}})} \preceq_{st} T_{(0,0,0,S_{\text{tot}})}$. It follows that increasing κ_5 will increase the mean first passage time from $(S_{\text{tot}}, 0, 0, 0)$ to $(0, 0, 0, S_{\text{tot}})$ when $\kappa_2 > \kappa_4$ (See Figure S.1). Indeed, when $\kappa_2 > \kappa_4$, it takes a longer time to get to $(0, 0, 0, S_{\text{tot}})$ from $(S_{\text{tot}}, 0, 0, 0)$ if reaction (5) is added to the system without that reaction.

On the other hand, suppose

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & -1 & 0 \end{bmatrix}$$

and infinitesimal transition rates $\check{\Upsilon}_1(x), \check{\Upsilon}_2(x), \check{\Upsilon}_3(x), \check{\Upsilon}_4(x)$ and $\check{\Upsilon}_5(x)$ are defined as for $\Upsilon_1(x), \Upsilon_2(x), \Upsilon_3(x), \Upsilon_4(x)$ and $\Upsilon_5(x)$ in (4.7), but with $\check{\kappa}_i = \kappa_i$, for $i = 1, 2, 3, 4$, $\check{\kappa}_5 \geq \kappa_5$, and $\kappa_2 < \kappa_4$. We can verify that the assumptions of Theorem S.2 hold, as follows. Condition (i) holds since $Av_1 = (1, -1, -1)^T$, $Av_2 = (0, 1, 1)^T$, $Av_3 = (1, 0, -1)^T$, $Av_4 = (0, 0, 1)^T$ and $Av_5 = (0, 1, 0)^T$. Assumption S.1 holds with $G^1 = \{1, 3\}$, $G^2 = \{2, 4\}$,

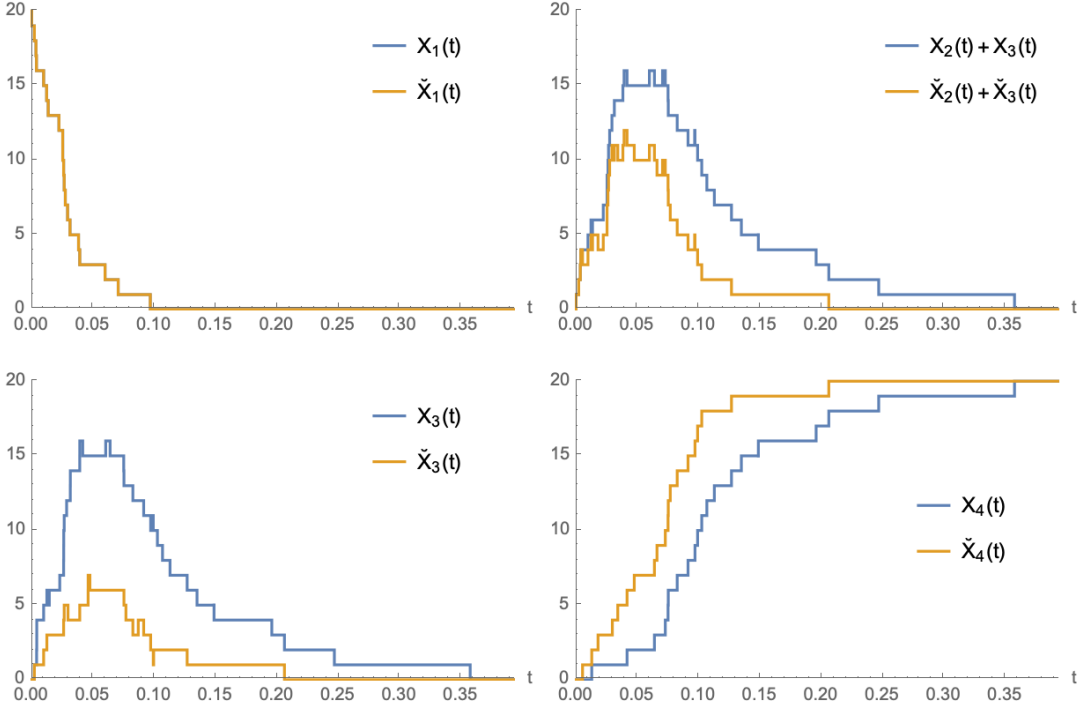


Figure S.1: **A typical coupled realization of sample paths for Example 4.3.** Here $S_{\text{tot}} = 20$, $\kappa_1 = \check{\kappa}_1 = 30$, $\kappa_2 = \check{\kappa}_2 = 50$, $\kappa_3 = \check{\kappa}_3 = 10$, $\kappa_4 = \check{\kappa}_4 = 10$, $\kappa_5 = 1000$ and $\check{\kappa}_5 = 10$. Both processes X and $\check{X}$ start at $(S_{\text{tot}}, 0, 0, 0)$. As shown in Section S.3.2, for these parameters, we have, almost surely, $X(t) \preccurlyeq_A \check{X}(t)$ for every $t \geq 0$ where the matrix A is given in (S.21). A coupled realization using the algorithm described in Section S.4 is plotted to illustrate this result. In particular, we see in this sample that $X_1(t) \geq \check{X}_1(t)$, $X_2(t) + X_3(t) \geq \check{X}_2(t) + \check{X}_3(t)$ and $X_3(t) \geq \check{X}_3(t)$ for all times t . Moreover, the first passage time to $(0, 0, 0, S_{\text{tot}})$ for X (which is equivalent to the first time to get to the state where $X_4 = S_{\text{tot}}$) is larger than for $\check{X}$. Since this first passage time result is true for all coupled samples, we can conclude that the mean first passage time from $(S_{\text{tot}}, 0, 0, 0)$ to $(0, 0, 0, S_{\text{tot}})$ is larger for X than for $\check{X}$.

$G^3 = \{5\}$ and $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$, $\sigma(4) = 4$, $\sigma(5) = 5$. To verify that condition (ii) of Theorem S.2 holds, let $x \in \mathcal{X}$, and first consider $y \in \partial_1(K_A + x) \cap \mathcal{X}$, where $\partial_1(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 = w_1, x_2 \geq w_2, x_2 + x_3 \geq w_2 + w_3, x_4 \leq w_4\}$. Given that $\langle A_{1\bullet}, v_1 \rangle = \langle A_{1\bullet}, v_3 \rangle = 1$, we need to check that $\Upsilon_1(x) + \Upsilon_3(x) \leq \check{\Upsilon}_1(y) + \check{\Upsilon}_3(y)$. Since $y \in \partial_1(K_A + x) \cap \mathcal{X}$, then $\Upsilon_1(x) = \kappa_1 x_1 = \kappa_1 y_1 = \check{\kappa}_1 y_1 = \check{\Upsilon}_1(y)$ and $\Upsilon_3(x) = \kappa_3 x_1 = \kappa_3 y_1 = \check{\kappa}_3 y_1 = \check{\Upsilon}_3(y)$, and so the desired inequality holds with equality. Secondly, consider $y \in \partial_2(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 \geq w_1, x_2 = w_2, x_2 + x_3 \geq w_2 + w_3, x_4 \leq w_4\}$. Given that $\langle A_{2\bullet}, v_2 \rangle = \langle A_{2\bullet}, v_5 \rangle = 1$ and $\langle A_{2\bullet}, v_1 \rangle = -1$, we need to check that $\Upsilon_1(x) \geq \check{\Upsilon}_1(y)$, $\Upsilon_2(x) \leq \check{\Upsilon}_2(y)$ and $\Upsilon_5(x) \leq \check{\Upsilon}_5(y)$. Since $y \in \partial_2(K_A + x) \cap \mathcal{X}$, then $\Upsilon_1(x) = \kappa_1 x_1 \geq \kappa_1 y_1 = \check{\kappa}_1 y_1 = \check{\Upsilon}_1(y)$, $\Upsilon_2(x) = \kappa_2 x_2 = \kappa_2 y_2 = \check{\kappa}_2 y_2 = \check{\Upsilon}_2(y)$ and $\Upsilon_5(x) = \kappa_5 x_2 = \kappa_5 y_2 \leq \check{\kappa}_5 y_2 = \check{\Upsilon}_5(y)$, and so the desired inequality holds. Lastly, consider $y \in \partial_3(K_A + x) \cap \mathcal{X} = \{w \in \mathcal{X} \mid x_1 \geq w_1, x_2 \geq w_2, x_2 + x_3 = w_2 + w_3, x_4 \leq w_4\}$. Given that $\langle A_{1\bullet}, v_1 \rangle = \langle A_{1\bullet}, v_3 \rangle = -1$ and $\langle A_{1\bullet}, v_2 \rangle = \langle A_{1\bullet}, v_4 \rangle = 1$, we need to check that $\Upsilon_1(x) + \Upsilon_3(x) \geq \check{\Upsilon}_1(y) + \check{\Upsilon}_3(y)$ and $\Upsilon_2(x) + \Upsilon_4(x) \leq \check{\Upsilon}_2(y) + \check{\Upsilon}_4(y)$. For $y \in \partial_3(K_A + x) \cap \mathcal{X}$, since $\kappa_2 < \kappa_4$, we have that $\Upsilon_1(x) + \Upsilon_3(x) = (\kappa_1 + \kappa_3)x_1 \geq (\kappa_1 + \kappa_3)y_1 = (\check{\kappa}_1 + \check{\kappa}_3)y_1 = \check{\Upsilon}_1(y) + \check{\Upsilon}_3(y)$ and

$\Upsilon_2(x) + \Upsilon_4(x) = \kappa_2 x_2 + \kappa_4 x_3 = (\kappa_4 - \kappa_2)x_3 + \kappa_2(x_2 + x_3) \leq (\kappa_4 - \kappa_2)y_3 + \kappa_2(y_2 + y_3) = \check{\kappa}_2 y_2 + \check{\kappa}_4 y_3 = \check{\Upsilon}_2(y) + \check{\Upsilon}_4(y)$. Thus, the conditions of Theorem S.2 are satisfied. In particular, we can conclude when $\kappa_2 < \kappa_4$, that it takes less time to get to $(0, 0, 0, S_{\text{tot}})$ from $(S_{\text{tot}}, 0, 0, 0)$ if reaction (5) is added to the system without that reaction⁹.

S.4 An algorithm for coupled stochastic simulation

We now provide an algorithm for stochastic simulation of the coupled continuous-time Markov chains X and $\check{X}$ under the conditions of Theorems 3.1, 3.2, 3.3 or S.2, when the transitions rates are bounded on the state space, i.e., when (5.1) holds.

Algorithm: Stochastic simulation for coupled continuous-time Markov chains X and $\check{X}$.

Data: Integer $n \geq 1$, real $T > 0$, set $\mathcal{X} \subseteq \mathbb{Z}_+^d$, vectors $v_1, \dots, v_n$ in $\mathbb{Z}^d \setminus \{0\}$, $x^\circ, \check{x}^\circ$ in $\mathcal{X}$, functions $\Upsilon = (\Upsilon_1, \dots, \Upsilon_n)$ and $\check{\Upsilon} = (\check{\Upsilon}_1, \dots, \check{\Upsilon}_n)$ and integer $a \in \{1, 2, 3, 4\}$ to indicate which theorem is invoked (3.1, 3.2, 3.3 or S.2).

Result: Sample of initial time and subsequent potential jump times $T_0, T_1, \dots, T_N$ and associated states $X(T_0), X(T_1), \dots, X(T_N)$ and $\check{X}(T_0), \check{X}(T_1), \dots, \check{X}(T_N)$ for the continuous-time Markov chains X and $\check{X}$ in the time interval $[0, T]$.

$\lambda \leftarrow 1 + n \max \left\{ \sup_{x \in \mathcal{X}} \sum_{j=1}^n \Upsilon_j(x), \sup_{x \in \mathcal{X}} \sum_{j=1}^n \check{\Upsilon}_j(x) \right\};$

$K \leftarrow 0;$

$T_0 \leftarrow 0;$

while $T_K \leq T$ **do**

$T_{K+1} \leftarrow T_K + \text{Exponential}(\lambda);$

$K \leftarrow K + 1;$

end

$Y_0, \check{Y}_0 \leftarrow x^\circ, \check{x}^\circ;$

$X(T_0), \check{X}(T_0) \leftarrow x^\circ, \check{x}^\circ;$

$N \leftarrow K - 1;$

if $N \geq 1$ **then**

for $k \leftarrow 0$ **to** $N - 1$ **do**

$U \leftarrow \text{Uniform}([0, 1]);$

$Y_{k+1}, \check{Y}_{k+1} \leftarrow \text{TransitionDTMC}(\Upsilon, \lambda, Y_k, U, a),$

$\text{TransitionDTMC}(\check{\Upsilon}, \lambda, \check{Y}_k, U, a);$

$X(T_{k+1}), \check{X}(T_{k+1}) \leftarrow Y_{k+1}, \check{Y}_{k+1};$

end

end

The random variables $T_1, \dots, T_N$ are called potential jump times because it could be that $X(T_k) = X(T_{k+1})$ or $\check{X}(T_k) = \check{X}(T_{k+1})$ for some $0 \leq k \leq N - 1$. Letting $T_{N+1} := T$, the trajectories of X are given by $X(t) = X(T_k)$ for $T_k \leq t < T_{k+1}$, $0 \leq k \leq N$, and similarly for the trajectories of $\check{X}$.

⁹The system without reaction (5) can be obtained from the system with reaction (5) by setting $\kappa_5 = 0$. While strictly speaking a zero rate constant is not within our definition of mass action kinetics, our theory does cover propensity functions with such a zero rate constant.

The function **TransitionDTMC** can be found below. This function is meant to replicate $\Phi_\lambda(x, u)$ in (5.4) for the case of Theorems 3.1 and 3.2, and $\Psi_\lambda(x, u)$ in (5.34) for the case of Theorems 3.3 and S.2.

Function **TransitionDTMC**($\Upsilon, \lambda, x, u, a$)

Data: Integer $n \geq 1$, set $\mathcal{X} \subseteq \mathbb{Z}_+^d$, vectors $v_1, \dots, v_n$ in $\mathbb{Z}^d \setminus \{0\}$.

Input: Function $\Upsilon = (\Upsilon_1, \dots, \Upsilon_n)$, $\lambda > 0, x \in \mathcal{X}, u \in [0, 1]$. Integer $a \in \{1, 2, 3, 4\}$ to indicate which theorem is invoked (3.1, 3.2, 3.3 or S.2). For the case of Theorem 3.3 or S.2, include partition $\{G^1, \dots, G^s\}$ and bijection σ .

Output: State $x + v \in \mathcal{X}$.

$v \leftarrow 0$;

if $a \in \{1, 2\}$ **then** // The case of Theorem 3.1 or 3.2.

for $j \leftarrow 1$ **to** n **do**

if $\frac{j-1}{n} \leq u < \frac{j-1}{n} + \frac{\Upsilon_j(x)}{\lambda}$ **then**

$v \leftarrow v_j$;

end

end

end

else // The case of Theorem 3.3 or S.2.

$p_0 \leftarrow 0$;

for $k \leftarrow 1$ **to** s **do**

$p_k \leftarrow p_{k-1} + |G^k|$;

for $q \leftarrow p_{k-1} + 1$ **to** p_k **do**

if $\frac{p_{k-1}}{n} + \sum_{\ell=p_{k-1}+1}^{q-1} \frac{\Upsilon_{\sigma(\ell)}(x)}{\lambda} \leq u < \frac{p_{k-1}}{n} + \sum_{\ell=p_{k-1}+1}^q \frac{\Upsilon_{\sigma(\ell)}(x)}{\lambda}$ **then**

$v \leftarrow v_{\sigma(q)}$;

end

end

end

end

return $x + v$

Remark S.3. The above algorithm can be adapted to provide simultaneous stochastic simulation for X and $\check{X}$ when transition rates are not bounded on the state space, by applying the algorithm on a sequence of bounded sets, which expand to the whole state space. This employs a sequence of successively defined stopping times $\tau_0 = 0$, $\tau_\ell = \inf\{t \geq \tau_{\ell-1} \mid X(t) \notin C_\ell \text{ or } \check{X}(t) \notin C_\ell\}$, $\ell = 1, 2, \dots$, where the C_ℓ are compact, $C_\ell \subseteq C_{\ell+1}$ for $\ell = 1, 2, \dots$ and $\cup_{\ell=1}^\infty C_\ell = \mathcal{X}$. The simulation of the pair $(X, \check{X})(t)$ for $\tau_\ell \leq t < \tau_{\ell+1}$ uses the above algorithm on C_ℓ for $\ell = 1, 2, \dots$.

References

- [1] MEYN, S. P., AND TWEEDIE, R. L.(1993). Stability of Markovian processes iii: Foster–Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548.
- [2] SHARPE, M. (1988). *General Theory of Markov Processes*. Academic Press.