

Individual Molecules Dynamics in Reaction Network Models

Original

Individual Molecules Dynamics in Reaction Network Models / Cappelletti, D; Rempala, Ga. - In: SIAM JOURNAL ON APPLIED DYNAMICAL SYSTEMS. - ISSN 1536-0040. - 22:2(2023), pp. 1344-1382. [10.1137/21M1459563]

Availability:

This version is available at: 11583/2982255 since: 2023-09-18T11:12:18Z

Publisher:

SIAM

Published

DOI:10.1137/21M1459563

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

1 **TRACKABLE SPECIES DYNAMICS IN REACTION NETWORK**
2 **MODELS***

3 DANIELE CAPPELLETTI[†] AND GRZEGORZ REMPALA[‡]

4 **Abstract.** In a stochastic reaction network setting we define a subset of species as 'trackable' if
5 we can consistently follow the fate of its individual molecules. We show that using the classical large
6 volume limit results, we may approximate the dynamics of a single molecule of trackable species in
7 a simple and computationally efficient way. We give examples on how this approach may be used
8 to obtain various characteristics of single-molecule dynamics (for instance, the distribution of the
9 number of infections in a single individual in the course of an epidemic or the activity time of a
10 single enzyme molecule). Moreover, we show how to approximate the overall dynamics of trackable
11 species in the full system with a collection of independent single-molecule trajectories, and give
12 explicit bounds for the approximation error in terms of the reaction rates. This approximation,
13 which is well defined for all times, leads to an efficient and fully parallelizable simulation technique
14 for which we provide some numerical examples.

15 **Key words.** Single-molecule dynamics, mathematical epidemiology, law of large numbers, Pois-
16 son process representation, stochastic approximation, dynamic survival analysis, Skorohod topology

17 **AMS subject classifications.** 60J28, 92C40, 92C42, 60F05

18 **1. Introduction.** Recent advances in modeling molecular systems, especially
19 our improved ability to track individual proteins, and the deluge of data from the
20 observations of both molecular and macro system (think, for instance, of the ongoing
21 COVID-19 pandemic), have created new scientific challenges of considering models
22 of very high resolution where the dynamics of a specific bio-molecule or a particu-
23 lar individual are of interest. In general, such 'agent-based' models are known to be
24 computationally very costly, due to complex stochastic dynamics and highly noisy
25 behavior of individual agents. However, it appears that, at least in some cases, simple
26 yet satisfactory approximation of individual molecular trajectory may be directly
27 inferred with the help of a classical approach of stochastic chemical kinetics that as-
28 sumes that all molecules or individuals are indistinguishable and consequently focuses
29 only on their aggregated counts. As an example of one such idea, originally proposed
30 in [7] and latter expanded in [15], consider the stochastic 'susceptible-infected' (*SI*)
31 chemical reaction network where a collection of $m + n$ molecules (or individuals) is
32 partitioned into two types: susceptible (*S*) and infected (*I*) with initially n being of
33 type *S* and remaining m of type *I*. The stochastic network evolves in time according
34 to a Markov jump process that counts the 'infection events', that is, the interactions
35 of one molecule of *I*-type with one molecule of *S*-type. Each such interaction creates
36 a new molecule of *I*-type and removes one of *S*-type (equivalently, a molecule changes
37 its type from *S* to *I*). Accordingly, in the reaction network notation described below
38 in Section 2.2 this model may be represented as



40 If the rate constant of the above reaction is β/n and we assume the usual mass action
41 kinetics [6], it is well know that the above stochastic reaction network satisfies the

*Submitted to the editors DATE.

Funding: DC was supported by the MIUR grant Dipartimenti di Eccellenza 2018-2022 (E11G18000350001). GAR was supported by the National Sciences Foundation grant (DMS-1853587).

[†]Politecnico di Torino, Torino, Italy (daniele.cappelletti@polito.it).

[‡]The Ohio State University, Columbus, Ohio, USA (rempala.3@osu.edu).

42 law of large numbers (admits the fluid limit approximation), in the sense that as
 43 $m, n \rightarrow \infty$ and $m/n \rightarrow \rho > 0$ the surviving proportion s_t of the S -type molecules
 44 follows the logistic equation that may be written in the form

$$45 \quad (1.2) \quad -\dot{s}_t/s_t = \beta(1 + \rho - s_t) \quad s_t(0) = 1.$$

46 Consequently, for $t \geq 0$ we have

$$47 \quad (1.3) \quad s_t = \frac{1 + \rho}{1 + \rho \exp(\beta(1 + \rho)t)}.$$

48 Thus, from the viewpoint of a single, randomly selected S -type molecule, the quantity
 49 s_t defines a *survival function* describing the limiting *probability* of surviving beyond
 50 time $t > 0$. The formula (1.3) led to the method of approximating the distribution
 51 of surviving molecules of S dubbed 'dynamic survival analysis' (DSA) described in
 52 [15] and applied recently to epidemic modeling [8, 9, 14, 21, 23]. The idea is further
 53 illustrated in Figure 1 where the average of the Markov process (1.1) is compared to
 54 the average of independent realizations of single molecule dynamics (which may be
 55 efficiently calculated using modern parallel computing capabilities). Note (1.2) may
 56 be also interpreted as the equation for the *hazard function* associated with s_t . This
 57 fact has some relevance for statistical inference, and is further exploited, for instance,
 58 in [9, 15].

59 Beyond the simple SI example, the DSA approach has been applied (mostly in
 60 the context of epidemics) only to a handful of reaction networks representing the so-
 61 called one-directional transfer models [7]. In all such networks individual molecules
 62 can only change their state in an ordered way, hence previously visited states are no
 63 longer attainable (for instance in the SI model a molecule of S -type can only change
 64 into I -type, but not vice-versa).

65 In the current paper we formally expand the survival function approach for track-
 66 ing the fate of individual molecules to a much broader class of networks, including
 67 those where molecules can return to their previous stages. A simple example is ob-
 68 tained by augmenting the SI network with the additional reaction $I \rightarrow S$, leading
 69 to the so-called SIS model (which is of interest in epidemiology) discussed in more
 70 detail in Example 4.3 below. To establish our results for such networks, we explore a
 71 different representation of the DSA approximation, which does not explicitly involve
 72 the survival function. Continuing with the SI model example, denote by $Y^i(t)$ the
 73 binary variable that takes value 1 or 0 according to whether i -th molecule is of type
 74 S or I . The limit dynamics of an i -th individual molecule (initially of type S) is then
 75 given by

$$76 \quad Y^i(t) = 1 - N^i \left(\beta \int_0^t Y^i(u)(1 + \rho - s_u) du \right)$$

77 where N^i is the unit Poisson process tracking the transition of the i -th molecule
 78 from S -type to I -type. Note that the argument of N^i is the cumulative hazard
 79 corresponding to integral of the right-hand side of (1.2) (see [15]). Such Poisson
 80 process representation is of course completely equivalent to simply having the time
 81 of switching of the i -th molecule from S to I follow the survival function (1.3), but
 82 it allows for a description of more complex scenarios than one-directional transfer
 83 models. For example, we will prove below that the limit dynamics of a single molecule

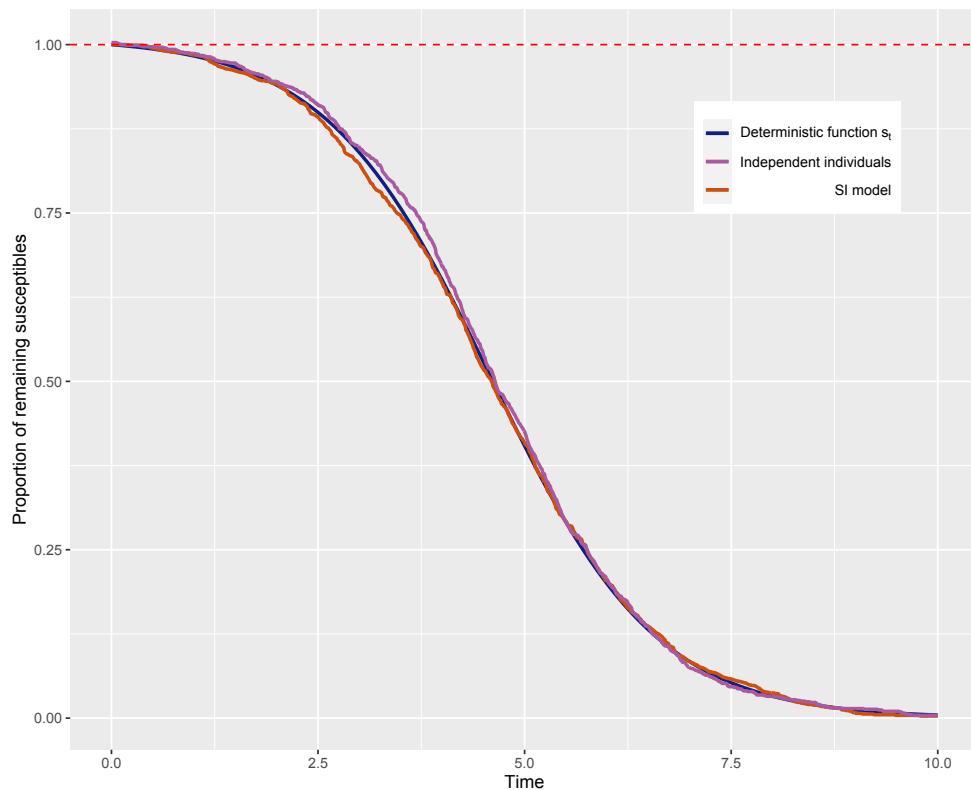


FIG. 1. **Survival approximation in the SI model.** The empirical trajectory of the proportion of the remaining S molecules in the SI model described in (1.1) as compared to the deterministic function s_t defined in (1.2) and the average of 1,000 independent single trajectories of individuals who become infected according to s_t . For the simulation we considered $n = 1,000$, $m = 10$, $\beta = 1$, and $\rho = 0.01$.

84 in the SIS model can be written as

$$85 \quad Y^i(t) = 1 - N_1^i \left(\beta \int_0^t Y^i(u)(1 + \rho - s_u) du \right) + N_2^i \left(\kappa \int_0^t (1 - Y^i(u)) du \right)$$

86 for independent and identically distributed unit-rate Poisson processes N_1^i and N_2^i .
87 Here, κ is the rate constant of the reaction $I \rightarrow S$.

88 In this work we study the Poisson process representation of the DSA approxima-
89 tion and give conditions under which it describes a single-molecule trajectory of the
90 original network. In particular, we explicitly derive error bounds of the DSA approxi-
91 mation, in terms of the underlying reaction network rates. We illustrate via numerical
92 examples how this novel technique could be useful to infer quantities pertaining to
93 single-molecule dynamics (such as the distribution of the number of infections a single
94 individual undergoes in a SIS model, or the time a single enzyme spends in the bound
95 state) in a computationally efficient way.

96 Further, we consider the problem of comparing the dynamics of an original full
97 reaction network with that of a collection of independent approximations of single-

98 molecule trajectories and provide explicit bounds on the error. Having the dynamics
 99 of the whole system approximated by a number of independent trajectories allows for
 100 computationally efficient simulation techniques, that are fully parallelizable. More-
 101 over, since the DSA approximation is defined for all times, it does not suffer from the
 102 problem of exiting the state space as it is known to happen in other methods such
 103 as diffusion approximations or tau leaping [4, 5, 12, 18]. Finally, the independence of
 104 the single-molecule trajectories also allows for much simplified statistical inferential
 105 procedures. Such applications were already considered in the context of SIR networks
 106 in recent papers on the Covid pandemic [14, 23]. A thorough investigation of these
 107 techniques in general reaction networks is currently being conducted and will appear
 108 in a future work.

109 The paper is organized as follows: in Section 2 we provide the necessary concepts
 110 pertaining to reaction network theory followed by the result on the approximation in
 111 classical scaling in Section 3. In Section 4 we give a formal definition of single-molecule
 112 trajectories of what we refer to as species that are 'trackable'. In Section 5 we state
 113 our main results. In particular, in Section 5.1 we give the theorem on the Poisson
 114 process representation of the DSA approximation for a single-molecule trajectory,
 115 and give examples of its applications in Section 5.2. Finally, in Section 5.3 we state
 116 the result on the approximation of the original full network via independent single-
 117 molecule trajectories, and give numerical examples. Proofs and explicit error bounds
 118 are given in the Appendix A.

119 2. Background definitions.

120 **2.1. Notation.** We denote by \mathbb{R} , $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$ the real, positive real, and non-
 121 negative real numbers, respectively. Similarly, we denote by \mathbb{Z} , $\mathbb{Z}_{\geq 1}$, and $\mathbb{Z}_{\geq 0}$ the real,
 122 positive real, and non-negative real numbers, respectively. Given a number $r \in \mathbb{R}$, we
 123 denote by $|r|$ its absolute value, and by $\lfloor r \rfloor$ the largest $m \in \mathbb{Z}$ such that $m \leq r$.

124 Given a vectors $v \in \mathbb{R}^n$, we denote its i th component by v_i , for all $1 \leq i \leq n$. We
 125 further denote

$$126 \quad \|v\|_{\infty} = \max_{1 \leq i \leq n} |v_i| \quad \text{and} \quad \lfloor v \rfloor = (\lfloor v_1 \rfloor, \dots, \lfloor v_n \rfloor).$$

127 Given two vectors $u, v \in \mathbb{R}_{\geq 0}^n$, we write

$$128 \quad u^v = \prod_{i=1}^m u_i^{v_i},$$

129 with the convention that $0^0 = 1$. We also write $u \geq v$ if the inequality holds
 130 component-wise. Furthermore, for any vector $v \in \mathbb{Z}_{\geq 0}^n$, we write

$$131 \quad v! = \prod_{i=1}^m v_i!.$$

132 Given a set A , we denote its cardinality by $\#A$ or, if it leads to no ambiguity, by $|A|$.
 133 We assume the reader is familiar with basic notions from stochastic process theory,
 134 such as the definition of continuous-time Markov chains and Poisson processes [19].

135 Consider a sequence of random variables $\{X_n\}_{n \in \mathbb{Z}_{\geq 0}}$ and a random variable X ,
 136 all defined on the same probability space and with values in a normed space $(E, \|\cdot\|)$.

137 We say that X_n converges in probability to X if for all $\eta \in \mathbb{R}_{>0}$

$$138 \quad \lim_{n \rightarrow \infty} P(\|X_n - X\| > \eta) = 0.$$

139 Given a topological space E we will denote by $D_E[0, T]$ the set of right-continuous
 140 left-bounded functions defined from $[0, T]$ to E , endowed with the Skorohod J_1 topol-
 141 ogy. In particular, we say that the sequence of processes $\{X_n\}$ with sample paths
 142 in $D_E[0, T]$ converges in probability to the process X (or simply that X_n converges
 143 in probability to X) if the Skorohod distance between X_n and X converges to 0 in
 144 probability (for more details, see for example [11, Chapter 3]).

145 **2.2. Stochastic reaction networks.** A *reaction network* is a triple $\mathcal{G} = \{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$ ■
 146 where (a) \mathcal{X} is an ordered finite sequence of d symbols, called *species*; (b) \mathcal{C} is a finite
 147 set of linear combinations of species over $\mathbb{Z}_{\geq 0}$, called *complexes*; (c) \mathcal{R} is a finite set
 148 of elements of $\mathcal{C} \times \mathcal{C}$, called *reactions*. We assume that no element of the form (y, y) is
 149 in \mathcal{R} , for any complex y , even though our results do not depend on this assumption.
 150 Following the usual notation of reaction network Theory, we further denote a reaction
 151 $(y, y') \in \mathcal{R}$ by $y \rightarrow y'$. We finally assume that each complex appears in at least
 152 one reaction, and that each species has a positive coefficient in at least one complex.
 153 Under this assumption and up to ordering of the set of species, a reaction network is
 154 uniquely determined by the set \mathcal{R} , or equivalently by the directed graph $(\mathcal{C}, \mathcal{R})$, called
 155 *reaction graph*. As an example, consider the reaction graph



157 In this case, the associated species are A , B , and C , $\mathcal{C} = \{A + B, 2B, B, C\}$, and
 158 $\mathcal{R} = \{A + B \rightarrow 2B, 2B \rightarrow A + B, B \rightarrow C\}$.

159 In this paper we will implicitly identify $\mathbb{R}^{|\mathcal{X}|}$ with \mathbb{R}^d , and therefore each $S \in \mathcal{X}$
 160 with a canonical basis vector of \mathbb{R}^d . With this in mind, the complexes are linear
 161 combination of species and can be therefore considered as vectors in $\mathbb{Z}_{\geq 0}^d$. As an
 162 example, if we order the species of (2.1) alphabetically, then the complex $A + B$ can be
 163 associated with the vector $(1, 1, 0)$, the complex $2B$ can be associated with $(0, 2, 0)$, the
 164 complex C with $(0, 0, 1)$, and so on. We will tacitly use the identification of complexes
 165 with integer vectors throughout the paper. Moreover, for each vector $v \in \mathbb{R}^d$ and for
 166 each species $S \in \mathcal{X}$ we denote by v_S the entry of v related to the canonical vector
 167 associated with S . We further define the *support* of v as $\text{supp}(v) = \{S \in \mathcal{X} : v_S > 0\}$.
 168 As an example, with the species of (2.1) alphabetically ordered, the support of $(1, 1, 0)$
 169 is $\{A, B\}$, the support of $(0, 2, 0)$ is $\{B\}$, and so on.

170 Deterministic and stochastic dynamical systems can be associated with a reaction
 171 network. The stochastic model is usually utilized when few individuals are present,
 172 so the stochastic component of the dynamic behaviour should not be ignored. In
 173 this case, the time evolution of the number of individuals of the different species is
 174 considered, for certain given propensities of the reactions to occur, and modeled via
 175 a continuous time Markov chain. More precisely, a *stochastic kinetics* for a reaction
 176 network \mathcal{G} is a correspondence between a reaction $y \rightarrow y'$ and a *rate function*
 177 $\lambda_{y \rightarrow y'} : \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, such that $\lambda_{y \rightarrow y'}(x) > 0$ only if $x \geq y$. A *stochastic reaction*
 178 *system* is a continuous time Markov chain $\{X(t) : t \geq 0\}$ with state space $\mathbb{Z}_{\geq 0}^d$ and
 179 transition rates from a state x to a state x' defined by

$$180 \quad q(x, x') = \sum_{\substack{y \rightarrow y' \in \mathcal{R} \\ y' - y = x' - x}} \lambda_{y \rightarrow y'}(x).$$

181 The associated generator is defined by

$$182 \quad Af(x) = \sum_{y \rightarrow y' \in \mathcal{R}} \lambda_{y \rightarrow y'}(x) (f(x + y' - y) - f(x))$$

183 for any function $f: \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}$ and any $x \in \mathbb{Z}_{\geq 0}^d$. Equivalently, the process X can be
184 described by

$$185 \quad X(t) = X(0) + \sum_{y \rightarrow y' \in \mathcal{R}} (y' - y) N_{y \rightarrow y'} \left(\int_0^\infty \lambda_{y \rightarrow y'}(X(s)) ds \right),$$

186 where the processes $\{N_{y \rightarrow y'}\}_{y \rightarrow y' \in \mathcal{R}}$ are independent unit-rate Poisson processes. For
187 more details on this representation, we refer to [6] or [11, Chapter 6].

188 In the deterministic setting, the concentration of the different species are assumed
189 to evolve according to an ordinary differential equation (ODE). Specifically, a *deter-*
190 *ministic kinetics* for a reaction network \mathcal{G} is a correspondence between the reactions
191 $y \rightarrow y'$ and the *rate function* $\lambda_{y \rightarrow y'}: \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, such that $\lambda_{y \rightarrow y'}(x) > 0$ only
192 if $x_i > 0$ whenever $y_i > 0$. A *deterministic reaction system* is the solution to the
193 ordinary differential equation

$$194 \quad (2.2) \quad \frac{d}{dt} Z(t) = \sum_{y \rightarrow y' \in \mathcal{R}} (y' - y) \lambda_{y \rightarrow y'}(x).$$

195 While our results hold in a more general scenario, all the simulations we show as-
196 sume *mass-action kinetics*, a popular choice of kinetics derived by the assumption that
197 all the reactants are well-mixed in the available volume [6]. Specifically, a stochastic
198 reaction system is a *stochastic mass-action system* if for every reaction $y \rightarrow y' \in \mathcal{R}$
199 we have

$$200 \quad \lambda_{y \rightarrow y'}(x) = \kappa_{y \rightarrow y'} \frac{x!}{(x - y)!} \mathbb{1}_{\{x \geq y\}},$$

201 for some positive constant $\kappa_{y \rightarrow y'}$ called *rate constant*. Similarly, a deterministic re-
202 action system is a *deterministic mass-action system* if for every reaction $y \rightarrow y' \in \mathcal{R}$
203 we have

$$204 \quad \lambda_{y \rightarrow y'}(x) = \kappa_{y \rightarrow y'} x^y,$$

205 for some positive constant $\kappa_{y \rightarrow y'}$ also called *rate constant*.

206 **3. Classical scaling.** Consider a reaction network $\mathcal{G} = \{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$, and a family
207 of stochastic kinetics $\{\lambda_{y \rightarrow y'}^V : y \rightarrow y' \in \mathcal{R}\}$ indexed by V . Let X^V denote the
208 associated continuous time Markov chain. V should be thought to as a parameter
209 expressing the volume, or the magnitude of the number of the present individuals.
210 Under the following technical but reasonable assumption the classical scaling of [11, 16]
211 holds:

212 *Assumption 3.1.* We assume that for any reaction $y \rightarrow y' \in \mathcal{R}$ there exists a
213 locally Lipschitz function $\lambda_{y \rightarrow y'}: \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^d$ such that for any compact set $K \subset \mathbb{R}_{\geq 0}^d$
214 we have

$$215 \quad \lim_{V \rightarrow \infty} \sup_{z \in K} \left| \frac{\lambda_{y \rightarrow y'}^V(\lfloor Vz \rfloor)}{V} - \lambda_{y \rightarrow y'}(z) \right| = 0.$$

216 THEOREM 3.2. Assume that Assumption 3.1 holds. Furthermore, assume that
 217 the random variables $X^V(0)/V$ converge in probability to a constant z^* as V goes to
 218 infinity. Finally, let $\{Z(t) : t \geq 0\}$ be the unique solution to (2.2) with $Z(0) = z^*$.
 219 Then, for any $\varepsilon > 0$ and any $T > 0$

$$220 \quad \lim_{V \rightarrow \infty} P \left(\sup_{t \in [0, T]} \left\| \frac{X^V(t)}{V} - Z(t) \right\|_{\infty} > \varepsilon \right) = 0.$$

221 Note that the distribution of the fate of a single molecule is not given, since the
 222 classical scaling concerns average dynamics. The goal of this paper is to address this
 223 issue, by providing a technique to simulate an approximation of the time evolution of
 224 a single observable species, as described in the next section.

225 **4. Trackable species.** We consider a special set of reactants, and assume that
 226 we can consistently follow the fate of a single molecule of these reactants through its
 227 different transformations, as for a single individual in the SI model. In general, the
 228 reactants whose dynamics we want to follow may be less than the chemical species
 229 listed in \mathcal{X} , or may be *portions* of them as in Example 4.5 below. To deal with this
 230 general setting, we introduce the set of *trackable species* as a set $\tilde{\mathcal{X}}$ of symbols endowed
 231 with a function $\tau: \tilde{\mathcal{X}} \rightarrow \mathcal{X} \cup \{0\}$. The different trackable species will identify the
 232 possible states of the molecules whose fate we want to follow. Every such state is taken
 233 by the reactant molecule when the molecule is (part of) one of the chemical species
 234 of \mathcal{X} . The function τ will link every trackable species with the corresponding species
 235 in \mathcal{X} . The number of trackable species defined in this way can be less than, equal to,
 236 or larger than the number of species. The set $\tilde{\mathcal{X}}$ needs to include the special state
 237 Δ to denote the potential degradation of the tracked molecule, and we set $\tau\Delta = 0$.
 238 To simplify the notation, for all $x, y \in \mathbb{Z}_{\geq 0}^d$ and $\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}$ we denote by $\theta_y(\tilde{S}, x)$
 239 the probability that a certain molecule of species $\tau(\tilde{S})$ is chosen if $y_{\tau(\tilde{S})}$ molecules are
 240 uniformly drawn out of $x_{\tau(\tilde{S})}$ molecules of $\tau(\tilde{S})$ available. Specifically,

$$241 \quad \theta_y(\tilde{S}, x) = \begin{cases} \frac{\binom{x_{\tau(\tilde{S})} - 1}{y_{\tau(\tilde{S})} - 1}}{\binom{x_{\tau(\tilde{S})}}{y_{\tau(\tilde{S})}}} = \frac{y_{\tau(\tilde{S})}}{x_{\tau(\tilde{S})}} & \text{if } x_{\tau(\tilde{S})} \geq y_{\tau(\tilde{S})} \geq 1 \\ 0 & \text{otherwise} \end{cases}.$$

242 For completeness, we define $\theta_y(\Delta, x) = 0$. Finally, note that in reactions such as
 243 $2A \rightarrow B + C$ we can imagine a molecule of A is transformed into a molecule of B , while
 244 the other molecule of A turns into a molecule of C . If we are tracking the fate of A
 245 molecules and the reaction $2A \rightarrow B + C$ occurs, it is reasonable to assume the molecule
 246 we are tracking has a 50% change of turning into a molecule of B , and a 50% change
 247 of becoming a molecule of C . We denote these probabilities with $p_{2A \rightarrow B+C}(A, B)$
 248 and $p_{2A \rightarrow B+C}(A, C)$, respectively, and in general allow for different value choices, as
 249 along as $p_{2A \rightarrow B+C}(A, B) + p_{2A \rightarrow B+C}(A, C) = 1$. The definition of *stochastic reaction*
 250 *system with trackable species* in the most general setting is below.

251 DEFINITION 4.1 (Stochastic reaction system with trackable species). Let $\mathcal{G} =$
 252 $\{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$ be a reaction network. Consider a family of stochastic kinetics $\{\lambda_{y \rightarrow y'}^V :$
 253 $y \rightarrow y' \in \mathcal{R}\}$ indexed by V , and let X^V denote the associated continuous time Markov
 254 chains. Let $\tilde{\mathcal{X}}$ be a set of trackable species. We define the stochastic reaction system
 255 with trackable species as the continuous-time Markov chain (Y^V, X^V) with state space

256 $\tilde{\mathcal{X}} \times \mathbb{Z}_{\geq 0}^d$ and transition rates

$$257 \quad q\left((\Delta, x), (\tilde{S}', x')\right) = \mathbb{1}_{\{\tilde{S}'\}}(\Delta) \sum_{\substack{y \rightarrow y' \in \mathcal{R} \\ y' - y = x' - x}} \lambda_{y \rightarrow y'}^V(x)$$

258
259 and for all $\tilde{S} \neq \Delta$

$$260 \quad q\left((\tilde{S}, x), (\tilde{S}', x')\right) = \sum_{\substack{y \rightarrow y' \in \mathcal{R} \\ y' - y = x' - x}} \left((1 - \theta_y(\tilde{S}, x)) \mathbb{1}_{\{\tilde{S} = \tilde{S}'\}} + \theta_y(\tilde{S}, x) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \right) \lambda_{y \rightarrow y'}^V(x),$$

261
262 where for all reactions $y \rightarrow y' \in \mathcal{R}$ the following holds:

- 263 • for any $\tilde{S} \in \tilde{\mathcal{X}}, \tilde{S}' \in \tilde{\mathcal{X}} \cup \{\Delta\}$ we have $0 \leq p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \leq 1$;
- 264 • $p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') = 0$ whenever $\tau(\tilde{S}) \notin \text{supp}(y)$ or $\tau(\tilde{S}') \notin \text{supp}(y') \cup \Delta$;
- 265 • if $\tau(\tilde{S}) \in \text{supp}(y)$ then

$$266 \quad \sum_{\tilde{S}' \in \tilde{\mathcal{X}} : \tau(\tilde{S}') \in \text{supp}(y') \cup \Delta} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') = 1.$$

267 In the above definition, the usual stochastic reaction system is coupled with the
 268 fate of a single trackable molecule: a trackable molecule in state \tilde{S} can transform
 269 whenever a reaction $y \rightarrow y'$ occurs, with a probability given by $\theta_y(\tilde{S}, \tilde{S}')$. By defini-
 270 tion, the quantity $\theta_y(\tilde{S}, \tilde{S}')$ denotes precisely the probability that the tracked molecule
 271 takes part in the reaction $y \rightarrow y'$, assuming that the reacting molecules are uniformly
 272 chosen among those present. If that happens, the new state of the tracked molecule is
 273 drawn according to the probability distribution $\{p_{y \rightarrow y'}(\tilde{S}, \tilde{S}')\}_{\tilde{S}' \in \text{supp}(y') \cup \Delta}$ (see Ex-
 274 ample 4.4 for a case where this distribution is non-trivial). If the tracked molecule is
 275 irreversibly degraded, its state becomes Δ and cannot be changed. In what follows,
 276 we will sometimes identify the state space of Y^V , given by $\tilde{\mathcal{X}}$, with the canonical basis
 277 of $\mathbb{R}^{|\tilde{\mathcal{X}}|}$, similarly to how complexes are implicitly identified with vectors in \mathbb{R}^d .

278 The only technical requirement to have trackable species is that in every reaction,
 279 every piece of the reactants on the left-hand side of the reaction either transforms
 280 into a piece of the reactants on the right-hand side, or is discarded. Mathematically,
 281 mapping the chemical species of the left-hand side of the reaction with those on the
 282 right-hand side is always possible by mapping to Δ potential species in excess on
 283 the left-hand side, so in principle the requirements of Definition 4.1 can always be
 284 satisfied with the choice $\tilde{\mathcal{X}} = \mathcal{X}$ and τ being the identity. Moreover, if we consider
 285 the physical system modelled by the reaction network, it is always true that reactions
 286 either transform molecules or degrade them. Hence, even when considering the phys-
 287 ical meaning of the model, trackable species can always be defined to track the fate of
 288 every molecule of particular interest. In this case however some care should be taken
 289 to reflect the real physical changes caused by the reactions, and the set $\tilde{\mathcal{X}}$ may need
 290 to be different from \mathcal{X} , as in Example 4.5.

291 *Remark 4.2.* The generator of a stochastic reaction system with trackable species,
 292 as defined in Definition 4.1, is given by

$$293 \quad \mathcal{A}f(\Delta, x) = \sum_{y \rightarrow y' \in \mathcal{R}} \lambda_{y \rightarrow y'}^V(x) \left(f(\Delta, x + y' - y) - f(\Delta, x) \right)$$

294 and for $\tilde{S} \neq \Delta$

295

$$296 \quad \mathcal{A}f(\tilde{S}, x) = \sum_{y \rightarrow y' \in \mathcal{R}} (1 - \theta_y(\tilde{S}, x)) \lambda_{y \rightarrow y'}^V(x) \left(f(\tilde{S}, x + y' - y) - f(\tilde{S}, x) \right)$$

$$297 \quad + \sum_{y \rightarrow y' \in \mathcal{R}} \sum_{\tilde{S}' \in \text{supp}(y') \cup \Delta} \theta_y(\tilde{S}, x) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \lambda_{y \rightarrow y'}^V(x) \left(f(\tilde{S}', x + y' - y) - f(\tilde{S}, x) \right),$$

298

299 for all functions $f: (\tilde{\mathcal{X}}) \times \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}$.

300 *Example 4.3.* Consider the SI reaction network described in (1.1), which we re-
 301 peat here for convenience:

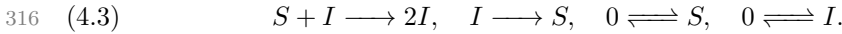


303 In this case, we are interested in describing the history of susceptible individuals
 304 who become infected. The set of trackable species is therefore $\tilde{\mathcal{X}} = \{\tilde{S}, \tilde{I}\}$ with
 305 $\tau(\tilde{S}) = S$ and $\tau(\tilde{I}) = I$. Furthermore, we choose the probabilities $p_{S+I \rightarrow 2I}(\tilde{S}, \tilde{I}) = 1$
 306 and $p_{S+I \rightarrow 2I}(\tilde{I}, \tilde{I}) = 1$. Alternatively, one can simply consider $\tilde{\mathcal{X}} = \{\tilde{S}\}$, with the
 307 understanding that whenever a susceptible individual gets infected we consider it as
 308 irreversibly degraded, and its state becomes Δ . In this case, $p_{S+I \rightarrow 2I}(\tilde{S}, \Delta) = 1$.

309 The state of single individuals can be tracked also in the more complex model

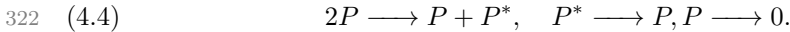


311 Here, the set of trackable species is $\{\tilde{S}, \tilde{I}\}$, with $\tau(\tilde{S}) = S$ and $\tau(\tilde{I}) = I$, and the trans-
 312 formation probabilities are $p_{S+I \rightarrow 2I}(\tilde{S}, \tilde{I}) = 1$, $p_{S+I \rightarrow 2I}(\tilde{I}, \tilde{I}) = 1$, $p_{I \rightarrow S}(\tilde{I}, \tilde{S}) = 1$.
 313 Here, relevant questions on the fate of a single individual could concern, for exam-
 314 ple, the number of infections it undergoes in a given time, or after how long the n th
 315 infection occurs. We can even extend the model to include migrations, and obtain



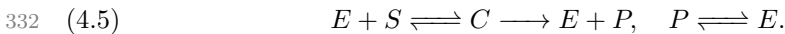
317 In this case, it is natural to assume $p_{S \rightarrow 0}(\tilde{S}, \Delta) = 1$ and $p_{I \rightarrow 0}(\tilde{I}, \Delta) = 1$. Relevant
 318 questions could involve, for example, the average number of infection a susceptible
 319 individual undergoes before migrating.

320 *Example 4.4.* Consider the following reaction network, where a protein P pro-
 321 motes its own phosphorylation:



323 Here, we may assume we are interested in observing the dynamics of a molecule
 324 of protein P . Hence, the set of trackable species is $\{\tilde{P}, \tilde{P}^*\}$ with $\tau(\tilde{P}) = P$ and
 325 $\tau(\tilde{P}^*) = P^*$. It is natural to assume that the two molecules of P involved in the
 326 reaction $2P \rightarrow P + P^*$ have the same probability of being phosphorylated or serving as
 327 the reaction catalyst. Hence, $p_{2P \rightarrow P+P^*}(\tilde{P}, \tilde{P}) = p_{2P \rightarrow P+P^*}(\tilde{P}, \tilde{P}^*) = 1/2$. The other
 328 transformation probabilities are given by $p_{P^* \rightarrow P}(\tilde{P}^*, \tilde{P}) = 1$ and $p_{P \rightarrow 0}(\tilde{P}, \Delta) = 1$.

329 *Example 4.5.* Consider the following reaction network, depicting a Michaelis-
 330 Menten mechanism where the product protein and the enzyme can spontaneously
 331 transform into each other:



333 In particular, the complex C represents a molecule of substrate S and enzyme bound
 334 together. When the bound is broken, it is natural to assume that the molecule of
 335 enzyme is released while the molecule of substrate is either released or transformed
 336 into the product P . Suppose we want to keep track of the history of a molecule of
 337 substrate S . If we were dealing with a classic Michaelis-Menten kinetics, i.e. without
 338 the reactions $P \rightleftharpoons E$, then we could simply consider as trackable species S , C , and P
 339 (or, more formally, S , C , and P with $\tau(\tilde{S}) = S$, $\tau(\tilde{C}) = C$, and $\tau(\tilde{P}) = P$). Indeed, a
 340 molecule of substrate can be either in free state (\tilde{S}), bound with the enzyme (\tilde{C}), or
 341 transformed into a product (\tilde{P}). Since the reactions $P \rightleftharpoons E$ are present, if we want
 342 to keep track of the fate of a molecule of substrate S we need to take into account
 343 the fact that it can transform into an enzyme, so E becomes a possible state of the
 344 molecule (more formally, \tilde{E} with $\tau(\tilde{E}) = E$). The problem now is that we need to
 345 differentiate between the portion of C that E and S get transformed into by the
 346 reaction $E + S \rightarrow C$: the portion of C that E gets transformed into will become a free
 347 enzyme again via the reaction $C \rightarrow E + P$, while the portion S gets transformed into
 348 will become a product. In order to formally express these dynamics, we consider as set
 349 of trackable species $\{\tilde{E}, \tilde{S}, \tilde{P}, \tilde{C}_E, \tilde{C}_S\}$, where \tilde{C}_E denotes we are tracking a molecule
 350 of E bound in the complex C , while \tilde{C}_S denotes we are tracking a molecule of S
 351 bound in C . The function τ associates every trackable species with its physical type:
 352 $\tau(\tilde{E}) = E$, $\tau(\tilde{S}) = S$, $\tau(\tilde{P}) = P$, $\tau(\tilde{C}_E) = C$, and $\tau(\tilde{C}_S) = C$. The transformation
 353 probabilities are given by

$$\begin{aligned}
 354 \quad p_{E+S \rightarrow C}(\tilde{E}, \tilde{C}_E) &= 1 & p_{C \rightarrow E+S}(\tilde{C}_E, \tilde{E}) &= 1 & p_{C \rightarrow E+P}(\tilde{C}_E, \tilde{E}) &= 1 \\
 p_{E+S \rightarrow C}(\tilde{S}, \tilde{C}_S) &= 1 & p_{C \rightarrow E+S}(\tilde{C}_S, \tilde{S}) &= 1 & p_{C \rightarrow E+P}(\tilde{C}_S, \tilde{P}) &= 1 \\
 p_{P \rightarrow E}(\tilde{P}, \tilde{E}) &= 1 & p_{E \rightarrow P}(\tilde{E}, \tilde{P}) &= 1 & &
 \end{aligned}$$

355 *Remark 4.6.* The interpretation of a stochastic reaction system with trackable
 356 species is that of a regular stochastic reaction system with the subsequent transfor-
 357 mations of a given particle being tracked. If the initial state $Y^V(0)$ of the tracked
 358 molecule is not present in the initial $X^V(0)$, that is if $X_{\tau(Y^V(0))}^V(0) = 0$, then the
 359 initial condition of (Y^V, X^V) is not consistent with the interpretation of the process.
 360 The process (Y^V, X^V) is still well-defined and its evolution can be studied, but its
 361 interpretation is no longer valid. In order to obtain meaningful results, we therefore
 362 tacitly assume that $X_{\tau(Y^V(0))}^V(0) > 0$, even if we do not require it formally.

363 **4.1. Representation as a regular stochastic reaction network.** In this
 364 section we show how a reaction network with trackable species (Y^V, X^V) can be
 365 realized as a regular stochastic reaction network with species set given by $\tilde{\mathcal{X}} \sqcup \mathcal{X}$,
 366 where \sqcup denotes a disjoint union. In particular, the state space is $\mathbb{Z}_{\geq 0}^{|\tilde{\mathcal{X}}|} \times \mathbb{Z}_{\geq 0}^d$, where
 367 for convenience we consider the first coordinates to refer to $\tilde{\mathcal{X}}$, and the rest to the
 368 species of the original process \mathcal{X} . We denote by (\tilde{x}, x) a generic state in $\mathbb{Z}_{\geq 0}^{|\tilde{\mathcal{X}}|} \times \mathbb{Z}_{\geq 0}^d$.
 369 Consider the set of reactions $\mathcal{R} \cup \tilde{\mathcal{R}}$ where

$$370 \quad \tilde{\mathcal{R}} = \{\tilde{S} + y \rightarrow \tilde{S}' + y' : y \rightarrow y' \in \mathcal{R}, \tilde{S}, \tilde{S}' \in \tilde{\mathcal{X}} \text{ and } p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') > 0\}$$

371 and endow them with the following reaction rates:

$$\begin{aligned}
 372 \quad \lambda_{y \rightarrow y'}^V(\tilde{x}, x) &= \sum_{\tilde{S} \in \tilde{\mathcal{X}}} \tilde{x}_{\tilde{S}} (1 - \theta_y(\tilde{S}, x)) \lambda_{y \rightarrow y'}^V(x) \\
 373 \quad \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\tilde{x}, x) &= \tilde{x}_{\tilde{S}} \theta_y(\tilde{S}, x) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \lambda_{y \rightarrow y'}^V(x). \\
 374
 \end{aligned}$$

375 Note that the second component of the process has the same transitions as X^V , with
 376 exactly the same rates. Hence, we can safely denote the process associated with the
 377 above stochastic reaction network by (\tilde{Y}^V, X^V) . Note that the quantity $\sum_{\tilde{s} \in \tilde{\mathcal{X}}} \tilde{x}_{\tilde{s}}$
 378 is conserved by all possible transitions. Hence, if we consider an initial condition
 379 $(\tilde{Y}(0), X(0))$ with $\sum_{\tilde{s} \in \tilde{\mathcal{X}}} \tilde{Y}_{\tilde{s}}(0) = 1$, then at any time point t exactly one entry of the
 380 vector $\tilde{Y}(t)$ is 1, and the other entries are zero. It follows that there is a bijection
 381 between the possible values of \tilde{Y} and $\tilde{\mathcal{X}}$, given by the function $\text{supp}(\tilde{Y}(t))$. In this
 382 case, by identifying trackable species with vectors of the canonical basis of $\mathbb{R}^{|\tilde{\mathcal{X}}|}$ as
 383 already done in the paper for the species in \mathcal{X} , the transition rates can be equivalently
 384 written as

$$385 \quad \lambda_{y \rightarrow y'}^V(\tilde{x}, x) = \sum_{\tilde{s} \in \tilde{\mathcal{X}}} \mathbb{1}_{\{\tilde{s}\}}(\tilde{x})(1 - \theta_y(\tilde{S}, x)) \lambda_{y \rightarrow y'}^V(x)$$

$$386 \quad \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\tilde{x}, x) = \mathbb{1}_{\{\tilde{s}\}}(\tilde{x}) \theta_y(\tilde{S}, x) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \lambda_{y \rightarrow y'}^V(x),$$

388 Hence, if $\sum_{\tilde{s} \in \tilde{\mathcal{X}}} \tilde{Y}_{\tilde{s}}(0) = 1$ then the transitions and the rates of (Y^V, X^V) and
 389 (\tilde{Y}^V, X^V) coincide, and (Y^V, X^V) can be therefore realized as a stochastic reaction
 390 network with an appropriate initial condition. In particular, we can write

(4.6)

$$391 \quad X^V(t) = X^V(0) + \sum_{y \rightarrow y' \in \mathcal{R}} (y' - y) N_{y \rightarrow y'} \left(\int_0^t \lambda_{y \rightarrow y'}^V(X^V(s)) ds \right)$$

(4.7)

$$392 \quad Y^V(t) = Y^V(0) + \sum_{y+\tilde{S} \rightarrow y'+\tilde{S}' \in \tilde{\mathcal{R}}} (\tilde{S}' - \tilde{S}) N_{y+\tilde{S} \rightarrow y'+\tilde{S}'} \left(\int_0^t \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(Y^V(s), X^V(s)) ds \right)$$

394 where N_r for $r \in \mathcal{R} \cup \tilde{\mathcal{R}}$ are independent unit-rate Poisson processes. Note that with
 395 the above writing, all the processes in the set $\{(Y^V, X^V)\}_{V \in \mathbb{Z}_{\geq 1}}$ can be defined on
 396 the same probability space.

397 **5. Results.** In this section we state the main results of the current paper and
 398 their applications.

399 **5.1. Classical scaling for the fate of a single molecule.** In this section we
 400 state a law of large number for the process Y^V . In order to do this, we consider a family
 401 of stochastic reaction systems with trackable species (Y^V, X^V) , with V varying in the
 402 integer numbers greater than one. We then assume that Assumption 3.1 is satisfied
 403 for some locally Lipschitz functions $\lambda_{y \rightarrow y'}$, and denote by Z the solution to (2.2).
 404 Hence, we know by Theorem 3.2 that $V^{-1}X^V$ will converge to Z path-wise with the
 405 uniform convergence topology over compact intervals of time, for V going to infinity.

406 In this section we express (Y^V, X^V) by means of independent unit-rate Poisson
 407 processes, as in (4.6) and (4.7). With the notation introduced in the previous section
 408 in mind, we have the following first technical result:

409 **LEMMA 5.1.** *Assume that Assumption 3.1 holds. Then, for any $\tilde{S}+y \rightarrow \tilde{S}'+y' \in$
 410 $\tilde{\mathcal{R}}$, any $w \in \tilde{\mathcal{X}}$, and any compact set $K \subset \mathbb{R}_{>0}^d$ we have*

$$411 \quad (5.1) \quad \lim_{V \rightarrow \infty} \sup_{z \in K} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(w, [Vz]) - \lambda_{y \rightarrow y'}(w, z) \right| = 0,$$

412 where the function $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'} : \tilde{\mathcal{X}} \times \mathbb{R}_{\geq 0}^d$ is defined as

$$413 \quad \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(w, z) = \mathbb{1}_{\{w\}}(\tilde{S}) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') y_{\tau(\tilde{S})} \frac{\lambda_{y \rightarrow y'}(z)}{z_{\tau(\tilde{S})}}$$

414 if both $z_{\tau(\tilde{S})}$ and $y_{\tau(\tilde{S})}$ are positive, and zero otherwise. Moreover, the function
415 $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}$ is locally Lipschitz if restricted to $\tilde{\mathcal{X}} \times \mathbb{R}_{> 0}^d$.

416 *Proof.* If $y_{\tau(S)} = 0$, then both $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V$ and $\lambda_{y \rightarrow y'}$ are constantly zero,
417 hence (5.1) holds. If $y_{\tau(S)}$ is positive, then for all $z \in K$ we have

$$418 \quad \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(w, [Vz]) - \lambda_{y \rightarrow y'}(w, z) \right| =$$

$$419 \quad \mathbb{1}_{\{w\}}(\tilde{S}) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \left| \theta_y(\tilde{S}, [Vz]) \lambda_{y \rightarrow y'}^V([Vz]) - y_{\tau(S)} \frac{\lambda_{y \rightarrow y'}(z)}{z_{\tau(S)}} \right|$$

422 Let $m = \min_{z \in K} z_{\tau(\tilde{S})}$, which is positive because K is a compact set contained in $\mathbb{R}_{> 0}^d$.
423 If V is large enough such that $Vm > y_{\tau(\tilde{S})}$ then

$$424 \quad \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(w, [Vz]) - \lambda_{y \rightarrow y'}(w, z) \right| =$$

$$425 \quad \mathbb{1}_{\{w\}}(\tilde{S}) p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') y_{\tau(S)} \left| \frac{\lambda_{y \rightarrow y'}^V([Vz])}{V \cdot (\lfloor Vz_{\tau(\tilde{S})} \rfloor / V)} - \frac{\lambda_{y \rightarrow y'}(z)}{z_{\tau(S)}} \right|$$

428 Hence, (5.1) follows from Assumption 3.1 and

$$429 \quad \max_{z \in K} \left| \frac{\lfloor Vz_{\tau(\tilde{S})} \rfloor}{V} - z_{\tau(\tilde{S})} \right| \leq \frac{1}{V}.$$

430 To conclude the proof, we only need to show that $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}$ restricted to $\tilde{\mathcal{X}} \times \mathbb{R}_{> 0}^d$ is
431 locally Lipschitz. However, this follows from it being the product (up to multiplication
432 by a constant) of the two locally Lipschitz functions $z \mapsto 1/z_{\tau(\tilde{S})}$ and $\lambda_{y \rightarrow y'}$. \square

433 The main goal of this section is to prove a classical scaling limit for a single-
434 molecule trajectory. To this aim, define the process Y by
(5.2)

$$435 \quad Y(t) = Y(0) + \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} (\tilde{S}' - \tilde{S}) N_{\tilde{S}+y \rightarrow \tilde{S}'+y'} \left(\int_0^t \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(Y(s), Z(s)) ds \right).$$

436 Then, the following result holds, where we implicitly identify the states of Y^V and Y
437 with the canonical basis of $\mathbb{R}^{|\tilde{\mathcal{X}}|}$. Note that the assumption that all the components
438 of the solution Z are strictly positive in the time interval $[0, T]$ is made, but this is
439 only a mild restriction to avoid unnecessary technicality, and is always verified under
440 mass-action kinetics as long as $Z(0) \in \mathbb{R}_{> 0}^d$ (see Remark 5.3). The proof of the result
441 is postponed to Appendix A, where more precise bounds are given.

442 **THEOREM 5.2.** *Assume that Assumption 3.1 holds. Furthermore, assume that*
443 *the random variables $X^V(0)/V$ converge in probability to some $z^* \in \mathbb{R}_{> 0}^d$ as V goes*
444 *to infinity, and let $Z(0) = z^*$. Assume that the solution Z to (2.2) with $Z(0) = z^*$*

445 exists over the interval $[0, T]$ and that

$$446 \quad m = \min_{\substack{i=1,2,\dots,d \\ u \in [0, T]}} Z_i(u) > 0.$$

447 Finally, assume that $Y^V(0) = Y(0)$ for all positive integers V . Then

$$448 \quad (5.3) \quad \lim_{V \rightarrow \infty} \sup_{t \in [0, T]} P(Y^V(t) \neq Y(t)) = \lim_{V \rightarrow \infty} \sup_{t \in [0, T]} E[\|Y^V(t) - Y(t)\|_\infty] = 0.$$

449 *Remark 5.3.* If we consider mass-action kinetics, then the deterministic solutions
 450 never touch the boundaries, provided that the initial condition is strictly positive [22].
 451 In this case, the existence of m as assumed in Theorem 5.2 is then guaranteed by
 452 $z^* \in \mathbb{R}_{>0}^d$.

453 *Remark 5.4.* Theorem 5.2 implies finite dimensional distribution convergence of
 454 Y^V to Y in the following sense: for all $0 \leq t_1 < t_2 < \dots < t_n \leq T$ we have

$$455 \quad P\left(\max_{1 \leq i \leq n} \|Y^V(t_i) - Y(t_i)\|_\infty > 0\right) \leq \sum_{i=1}^n P(\|Y^V(t_i) - Y(t_i)\|_\infty > 0),$$

456 and the latter tends to 0 as V tends to ∞ , under the conditions of Theorem 5.2.

457 Some simulations of the process Y are proposed in Figure 2 for the case of the
 458 SIS model (4.2). We conclude this section with the following result, concerning the
 459 convergence of Y^V to Y as processes with sample paths in $D_{\tilde{\mathcal{X}}}[0, T]$. We note how
 460 this result is necessary for the convergence of continuous functionals of $D_{\tilde{\mathcal{X}}}[0, T]$, as
 461 highlighted in Section 5.2.

462 **THEOREM 5.5.** *Assume that Assumption 3.1 holds. Furthermore, assume that the*
 463 *random variables $X^V(0)/V$ converge weakly to a constant z^* as V goes to infinity,*
 464 *and let $Z(0) = z^*$. Assume that the solution Z to (2.2) with $Z(0) = z^*$ exists over*
 465 *the interval $[0, T]$ and that*

$$466 \quad m = \min_{\substack{S \in \tilde{\mathcal{X}} \\ u \in [0, T]}} Z_S(u) > 0.$$

467 Finally, assume that $Y^V(0) = Y(0)$ for all positive integers V . Then Y^V converges
 468 in probability to Y as processes with sample paths in $D_{\tilde{\mathcal{X}}}[0, T]$ (where we identify $\tilde{\mathcal{X}}$
 469 with the elements of the canonical basis of $\mathbb{R}^{|\tilde{\mathcal{X}}|}$ and embed it with the metric $\|\cdot\|_\infty$,
 470 or any equivalent one).

471 The proof is given in Appendix A.

472 **5.2. Applications of Theorem 5.5.** The convergence of Theorem 5.5 allows
 473 us to state convergence in probability of $f(Y^V)$ to $f(Y)$, where $f: D_{\tilde{\mathcal{X}}}[0, T] \rightarrow \mathbb{R}$ is
 474 a functional that is continuous with respect to the Skorohod J_1 topology. Classical
 475 examples are $f(x) = \sup_{t \in [0, T]} \|x(t)\|_\infty$, $f(x) = \int_0^T \phi(x(s)) ds$ for some continuous
 476 function ϕ , or $f(x) = \sup_{t \in [0, T]} (x(t) - x(t-))$ where $x(t-) = \lim_{h \uparrow t} x(h)$ (see for
 477 example [11, Chapter 3]). More concretely, a functional we may want to consider is
 478 the number of times an individual gets infected in the interval $[0, T]$, assuming the
 479 model of equation (4.2) is in place. We denote this functional by ψ . Note that the
 480 convergence of X^V/V to its deterministic fluid limit, as stated in Theorem 3.2, does
 481 not give any mean of inferring the distribution of $\psi(Y^V)$. However, knowing that

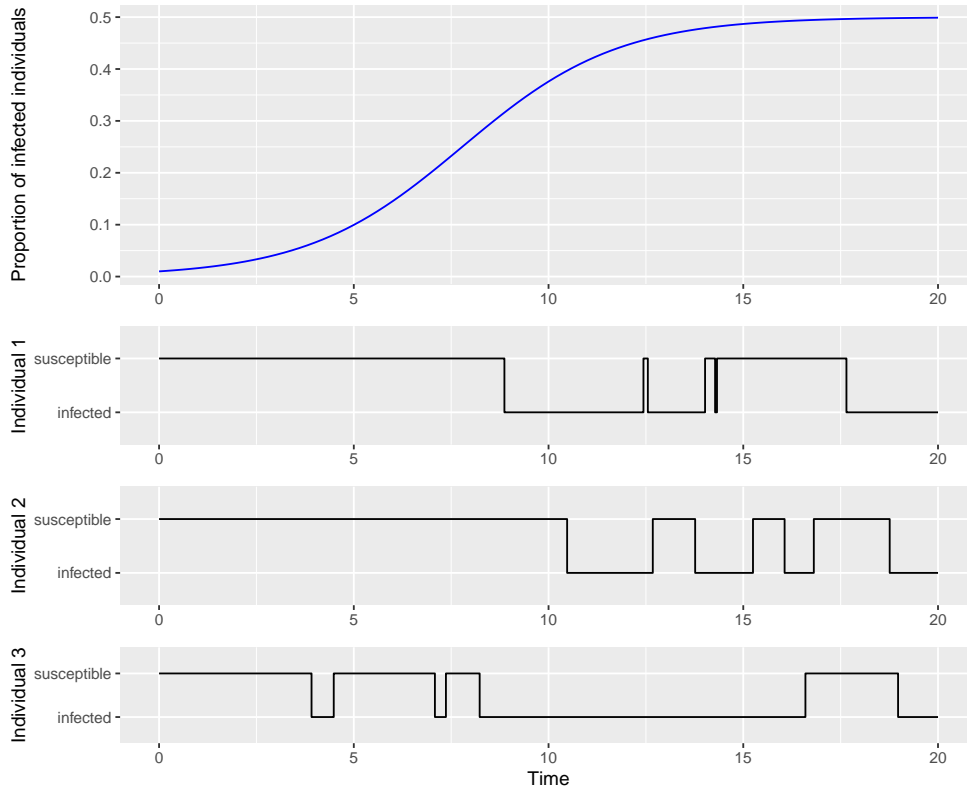
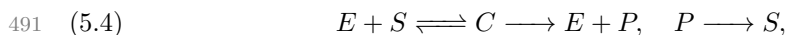


FIG. 2. *The process Y in SIS model.* Consider the model (4.2), and let Y be as in (5.2). The first panel shows the concentration of infected individuals Z_I according to the deterministic solution to (2.2) with $Z_S(0) = 0.99$ and $Z_I(0) = 0.01$. Mass-action kinetics is assumed, with the rate constants of $S + I \rightarrow 2I$ and $I \rightarrow S$ being 1 and 0.5, respectively. According to (5.2), Z_I determines the rate at which the single-individual process Y turns from 'susceptible' to 'infected'. The last three panels show independent realizations of Y . The times in the x-axes of the four panels are aligned.

482 $\psi(Y^V)$ converges in probability to $\psi(Y)$, if V is large enough we can approximate the
 483 distribution of the former by the distribution of the latter. Obtaining an estimate of
 484 the distribution of $\psi(Y)$ only requires the simulation of enough independent copies of
 485 Y , whose jump rates are deterministic and therefore do not require a simulation of
 486 X^V to be computed, as opposed to the much more expensive strategy of simulating
 487 multiple independent trajectories of (Y^V, X^V) via the Gillespie algorithm (which is
 488 especially cumbersome for large values of V). The empirical distributions obtained
 489 with the two strategies are compared in Figure 3. Similarly, we can apply our results
 490 to a Michaelis-Menten mechanism. Consider the model



492 where the enzyme activities counterbalances a spontaneous transformation of mole-
 493 cules of type P into molecules of type S . To measure the activity level of the enzymes,
 494 we may want to study for how long a randomly chosen enzyme molecule is in bound
 495 state C up to a given time T . Let us call this quantity $\nu(Y^V)$. The classical scal-

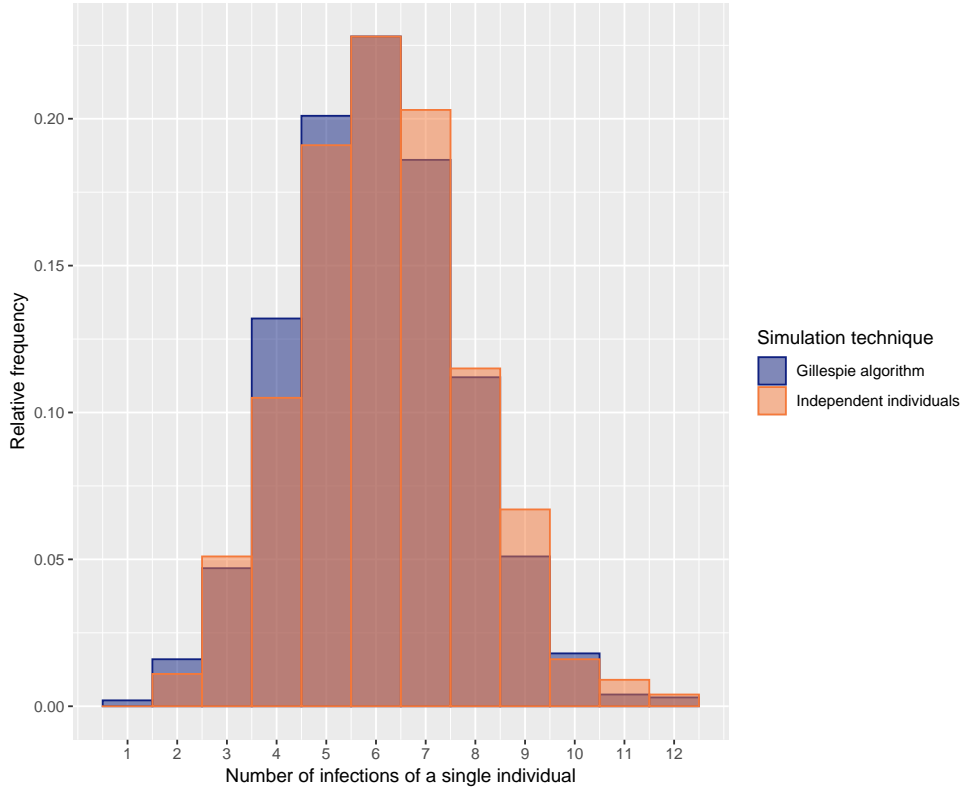


FIG. 3. **Empirical distribution of number of infections in SIS model.** Consider the model (4.2), and let ψ be the number of infections a randomly selected individual undergoes up to time T . The empirical distributions of $\psi(Y^V)$ and $\psi(Y)$ are compared, the former obtained by the simulation of 1,000 independent copies of (Y^V, X^V) via the Gillespie algorithm (applied to the formulation in terms of usual stochastic reaction networks discussed in Section 4.1), and the latter obtained via the simulation of 1,000 copies of Y . Here, $V = 1,000$ and the initial portion of infected individuals is 1% (so we are initially close to the boundary and we may expect some minor discrepancy between X^V/V and its deterministic limit Z , see also Figure 5). Mass-action kinetics is assumed, with the rate constants of $S + I \rightarrow 2I$ and $I \rightarrow S$ being 1 and 0.5, respectively.

496 ing of Theorem 3.2 does not allow for inference of the distribution of $v(Y^V)$, but
 497 Theorem 5.5 ensures that it converges to the distribution of $v(Y)$ as V tends to ∞ .
 498 Figure 4 compares the empirical distributions of $v(Y^V)$ and $v(Y)$ obtained by the
 499 simulation of 1,000 independent copies of (Y^V, X^V) and 1,000 independent copies of
 500 Y , respectively. For this comparison we chose $V = 1,000$.

501 **5.3. Approximating the system dynamics with single-molecule trajec-**
 502 **tories.** Let $\bar{\mathcal{X}} \subseteq \mathcal{X}$ be the set of species that can be tracked in some form:

$$503 \quad \bar{\mathcal{X}} = \{S \in \mathcal{X} : S = \tau(\tilde{S}) \text{ for some } \tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}\}.$$

504 Moreover, let $\pi: \mathbb{R}^d \rightarrow \mathbb{R}^{|\bar{\mathcal{X}}|}$ be the projection of the state space onto the coordinates
 505 relative to the species in $\bar{\mathcal{X}}$. The aim of this section is to approximate the dynamics
 506 of $\pi(X^V)$ by means of a sum of *independent* processes distributed as in (5.2) (po-
 507 tentially with rescaled dynamics, as shown in the statement of Theorem 5.8). Note

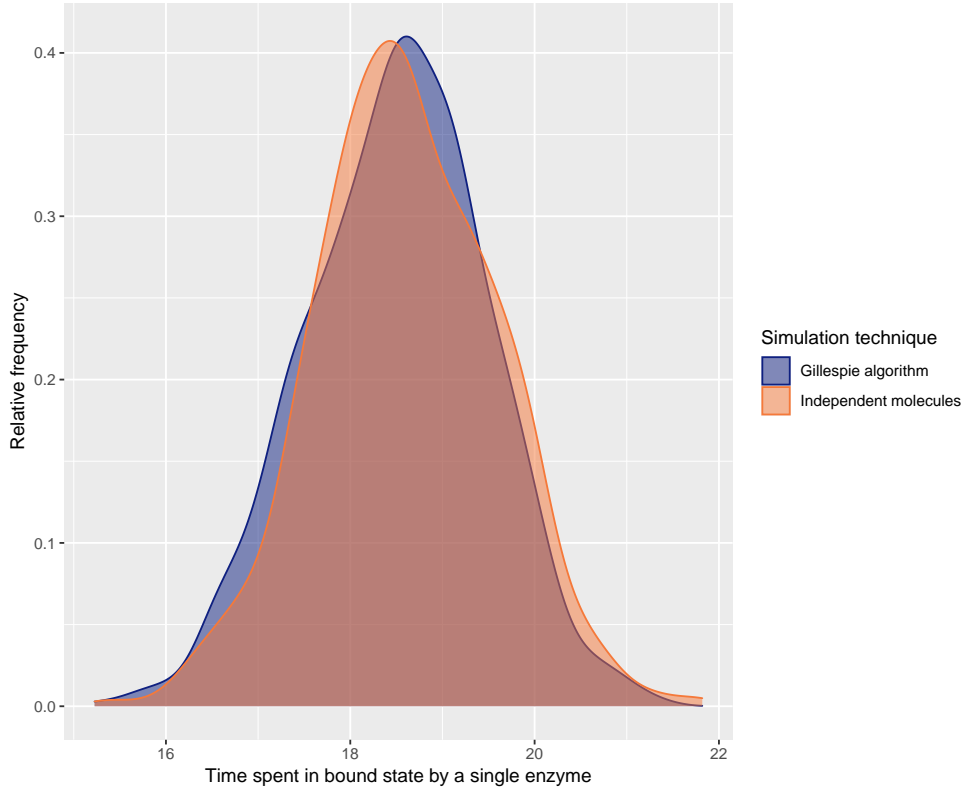


FIG. 4. **Empirical density of time in bound state in Michaelis-Menten model.** Consider the model (5.4), and let v be the time a randomly selected molecule of enzyme is in bound state C up to time T . The empirical distributions of $v(Y^V)$ and $v(Y)$ are compared, the former obtained by the simulation of 1,000 independent copies of (Y^V, X^V) via the Gillespie algorithm (applied to the formulation in terms of usual stochastic reaction networks discussed in Section 4.1), and the latter obtained via the simulation of 1,000 copies of Y . Here, $V = 1,000$ and $Z(0) = X(0)/V = (0.5, 10, 0.5, 1)$, where the species are ordered as in E, S, C, P . Mass-action kinetics is assumed, with the rate constants of $E + S \rightarrow C$, $C \rightarrow E + S$, $C \rightarrow E + P$, and $P \rightarrow S$ being 1, 5, 1, and 0.5, respectively.

508 that the goal of such an approximation is not to provide a faster simulation method
509 than those present in the literature: our goal is to break down the dynamics of sev-
510 eral correlated particles into a set of independent single-molecule trajectories which
511 could be simulated simultaneously by a highly parallelizable algorithm. We begin by
512 identifying each trackable species $\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}$ with a different physical portions of
513 the chemical species $\tau(\tilde{S})$: m molecules of species $S \in \mathcal{X}$ are available at time t if and
514 only if the quantity of each trackable species \tilde{S} satisfying $\tau(\tilde{S}) = S$ is m at time t .
515 Under this assumption, clearly the process X^V can be expressed in terms of the dy-
516 namics of its individual trackable species, which are typically not independent of each
517 other. We further restrict ourselves to models that are *sub-conservative* with respect
518 to the trackable species. This means that while trackable species can potentially be
519 degraded (by entering the fictitious state Δ), their total mass never increases. Equiv-
520 alently, we assume that each time a trackable species is created it is by transformation

521 of another trackable species. We assume sub-conservativeness because we want the
 522 single-molecule fates we track to be independent, while their agglomeration is still
 523 able to approximately describe the dynamics of the whole system. If we allowed for
 524 mass creation, we would need to introduce new molecules over time and track them.
 525 Defining the molecule creation times over a finite interval of time independently on
 526 each other is technically possible if the creation rate is deterministic: it is sufficient to
 527 first simulate a Poisson random variable counting the total number of new molecules
 528 in the finite time interval, then consider each creation time as independent on the
 529 others with probability density proportional to the deterministic creation rate. How-
 530 ever, this procedure requires the introduction of further notation and for the sake of
 531 clarity we decided to only present the simpler case of sub-conservative models (with
 532 respect to the trackable species).

533 *Assumption 5.6.* Let (Y^V, X^V) be a family of stochastic reaction systems with
 534 trackable species. We assume that for each reaction $y \rightarrow y' \in \mathcal{R}$ and for each $\tilde{S}' \in$
 535 $\tilde{\mathcal{X}} \setminus \{\Delta\}$

$$536 \quad \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') = y'_{\tau(\tilde{S}')}$$

537 For all $S \in \bar{\mathcal{X}}, \tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}$ define

$$538 \quad \tau^{-1}(S) = \{\tilde{S}' \in \tilde{\mathcal{X}} : \tau(\tilde{S}') = S\} \quad \text{and} \quad \alpha(S) = \#\tau^{-1}(S)$$

539 The sub-conservation of the model with respect to the trackable species is formally
 540 stated as follows.

541 *LEMMA 5.7.* Let (Y^V, X^V) be a family of stochastic reaction systems with track-
 542 able species satisfying Assumption 5.6. Then, for all $V \in \mathbb{Z}_{\geq 1}$ and for all $t \in \mathbb{R}_{>0}$

$$543 \quad (5.5) \quad \|\pi(X^V(t))\|_1 \leq \sum_{S \in \bar{\mathcal{X}}} \alpha(S) X_S^V(t) \leq \sum_{S \in \bar{\mathcal{X}}} \alpha(S) X_S^V(0).$$

544 *Proof.* The first inequality of (5.5) simply follows from the fact that the quantities
 545 $\alpha(S)$ are greater than or equal to 1. For the second inequality, simply note that if a
 546 reaction $y \rightarrow y' \in \mathcal{R}$ occurs at time t , then

$$\begin{aligned} 547 \quad \sum_{S \in \bar{\mathcal{X}}} \alpha(S) X_S^V(t) - \sum_{S \in \bar{\mathcal{X}}} \alpha(S) X_S^V(t-) &= \sum_{S \in \bar{\mathcal{X}}} \alpha(S) y'_S - \sum_{S \in \bar{\mathcal{X}}} \alpha(S) y_S \\ 548 \quad &= \sum_{\tilde{S}' \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y'_{\tau(\tilde{S}')} - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} \\ 549 \quad &= \sum_{\tilde{S}' \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} \\ 550 \quad &\leq \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})} = 0. \\ 551 \end{aligned}$$

552 Note that in the third equality we used Assumption 5.6, and in the last equality we
 553 used $\sum_{\tilde{S}' \in \tilde{\mathcal{X}} \setminus \{\Delta\}} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \leq 1$. Since the quantity $\sum_{S \in \bar{\mathcal{X}}} \alpha(S) X_S^V$ is not increasing
 554 with the occurrence of a reaction, (5.5) is proven. \square

555 The main result of this section is the following one, a more detailed version of
 556 which is proven in the Appendix. In particular, in Theorem A.5 a convergence rate
 557 of the order of $e^{-C\sqrt{V}}$ for a positive constant C is proven, provided that the initial
 558 conditions of X^V and \tilde{X}^V are close enough.

559 **THEOREM 5.8.** *Assume that Assumptions 3.1 and 5.6 are satisfied, and consider*
 560 *a family of stochastic reaction systems with trackable species (Y^V, X^V) . Assume that*
 561 *$V^{-1}X^V(0)$ converges in distribution to some $z^* \in \mathbb{R}_{>0}^d$ as V goes to infinity and*
 562 *$E[\pi(X^V(0))] < \infty$ for all $V \in \mathbb{Z}_{\geq 1}$. Let $\tilde{X}^V(0) = \lfloor Vz^* \rfloor$ and define the process \tilde{X}^V*
 563 *by*

$$564 \quad (5.6) \quad \tilde{X}^V(t) = \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \frac{\tau(Y^{\tilde{S},i}(t))}{\alpha(\tau(Y^{\tilde{S},i}(t)))},$$

565 where the processes $(Y^{\tilde{S},i})_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}, i \in \mathbb{Z}_{\geq 1}}$ are independent and satisfy

$$566 \quad Y^{\tilde{S},i}(t) = \tilde{S} + \sum_{\tilde{S}''+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} (\tilde{S}'' - \tilde{S}') N_{\tilde{S}''+y \rightarrow \tilde{S}''+y'}^{\tilde{S},i} \left(\int_0^t \lambda_{\tilde{S}''+y \rightarrow \tilde{S}''+y'}(Y^{\tilde{S},i}(u), Z(u)) du \right),$$

567 for a family of independent, identically distributed unit-rate Poisson processes $\{N_r^{\tilde{S},i}\}_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}, i \in \mathbb{Z}_{\geq 1}, r \in \tilde{\mathcal{R}}}$.
 568 Then,

$$569 \quad \lim_{V \rightarrow \infty} E \left[\sup_{0 \leq s \leq t} \left\| \frac{\pi(X^V(t))}{V} - \frac{\tilde{X}^V(t)}{V} \right\| \right] = 0.$$

570 Note that in the definition of \tilde{X}^V above we consider as many independent single-
 571 molecule trajectories as many trackable molecules are in the system at time 0. A
 572 natural question is whether a good approximation of the original model X^V can be
 573 obtained by considering the agglomeration of less independent single-molecule trajec-
 574 tories. However, a detailed study of the error in this case is out of the scope of the
 575 present paper.

576 *Example 5.9.* Consider the SIS model of equation (4.2). We assume $X_S^V(0) =$
 577 $0.99V$ and $X_I^V(0) = 0.01V$, and let $V = 1,000$. We wish to approximate the number
 578 of susceptible individuals by

$$579 \quad \frac{X_S^V(t)}{V} \approx \frac{\tilde{X}_S^V(t)}{V}.$$

580 In order to test the performance of the above approximation, we simulate 100 in-
 581 dependent copies of X^V and \tilde{X} , and plot them against each other in Figure 5. It
 582 is perhaps not surprising to note a higher variance for the trajectories of X^V with
 583 respect of those of \tilde{X}^V : the former is the result of several single-molecule trajec-
 584 tories that are naturally correlated with each other, specifically the rate at which a
 585 single molecule changes state is stochastic and given by the current state of all the
 586 other molecules. In the approximation, the dynamics of the single tracked molecules
 587 are independent and their rate of change from one state to another state are purely
 588 deterministic, which leads to fewer stochastic fluctuations. However, we do observe a
 589 discrepancy between the two models only at the beginning of the trajectories, when

590 the number of infected individuals is rather low (only 10 individuals in the initial con-
 591 dition) and the deterministic approximation given by Theorem 3.2 is perhaps not yet
 592 accurate enough. As a matter of fact, Figure 6 shows that the difference in variance
 593 is considerably reduced if the initial counts of infected individuals is increased to 100.
 594

595 We are interested in bounding

$$596 \quad (5.7) \quad P \left(\sup_{0 \leq t \leq T} \left| \frac{X_S^V(t)}{V} - \frac{\tilde{X}_S^V(t)}{V} \right| > \varepsilon \right),$$

597 for a fixed $\varepsilon \in \mathbb{R}_{>0}$. Assume mass-action kinetics and let κ_1 and κ_2 be the rate
 598 constants of $S + I \rightarrow 2I$ and $I \rightarrow S$, respectively. Moreover, assume for simplicity
 599 that $X^V(0) = \tilde{X}^V(0) = VZ(0)$ and $X_S^V(0) + X_I^V(0) = V$. Since the total number
 600 of individual is conserved, for all $0 \leq t \leq T$ we have $X_S^V(t) + X_I^V(t) = V$. By
 601 superposition there exist two independent unit-rate Poisson processes $\tilde{N}_{S+I \rightarrow 2I}$ and
 602 $\tilde{N}_{I \rightarrow S}$ such that for all $0 \leq t \leq T$ and for a fixed V we have (with a simplified notation
 603 that does not take into account the initial values of the independent single individual

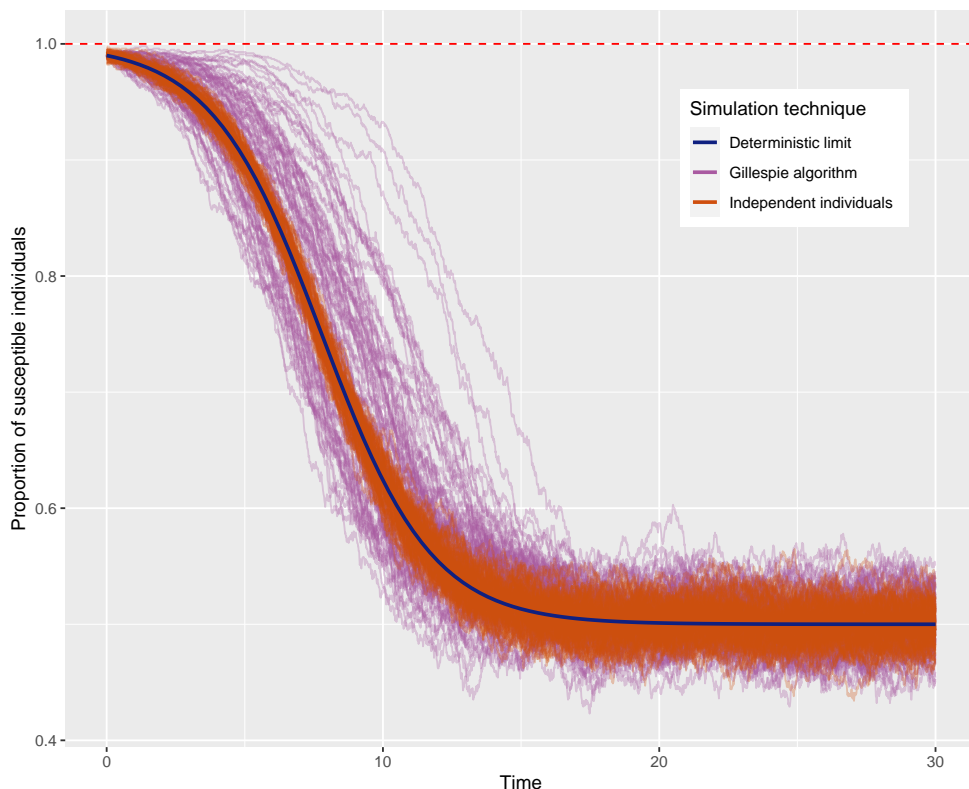


FIG. 5. *Comparison in SIS model.* Comparison of 100 independent trajectories of X_S^V/V and \tilde{X}_S^V/V , considering the SIS model described in (4.2). Here, $X_S^V(0) = 0.99V$, $X_I^V(0) = 0.01V$, and $V = 1,000$. Mass-action kinetics is assumed, with the rate constants of $S + I \rightarrow 2I$ and $I \rightarrow S$ being 1 and 0.5, respectively.

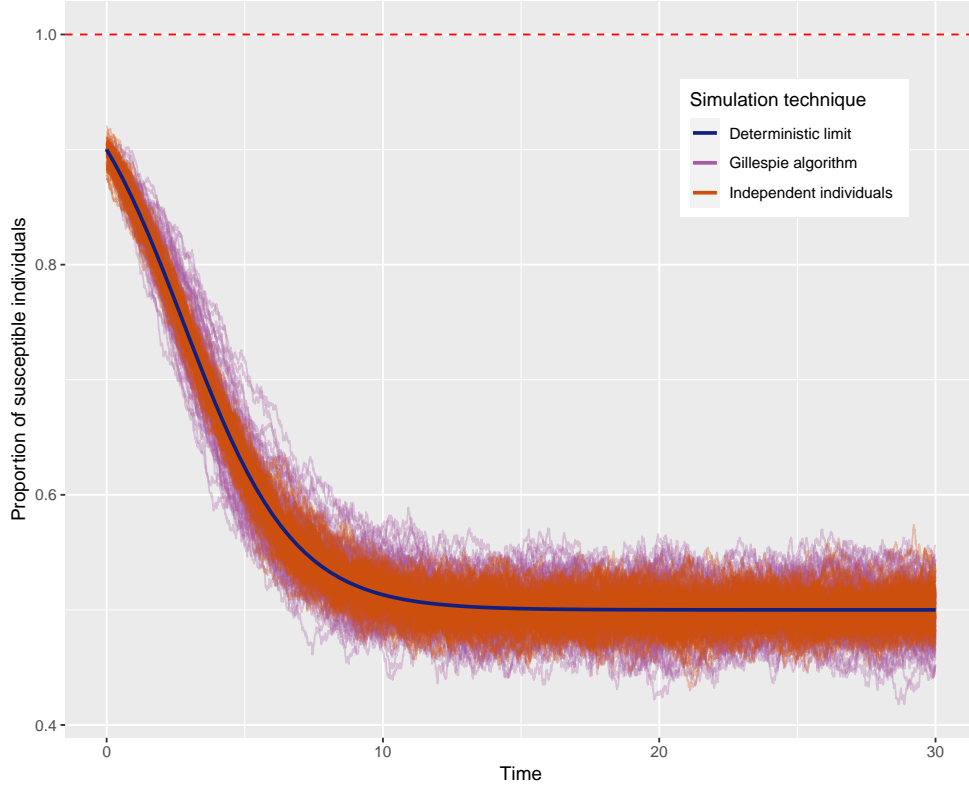


FIG. 6. *Comparison in SIS model.* Comparison of 100 independent trajectories of X_S^V/V and \tilde{X}_S^V/V , considering the SIS model described in (4.2). Here, $X_S^V(0) = 0.9V$, $X_I^V(0) = 0.1V$, and $V = 1,000$. Mass-action kinetics is assumed, with the rate constants of $S + I \rightarrow 2I$ and $I \rightarrow S$ being 1 and 0.5, respectively.

604 trajectories)

$$605 \quad \tilde{N}_{S+I \rightarrow 2I} \left(\int_0^t \kappa_1 \tilde{X}_S^V(u) Z_I(u) du \right) = \sum_{i=1}^V N_{\tilde{S}+S+I \rightarrow \tilde{I}+2I}^i \left(\int_0^t \mathbb{1}_{\{\tilde{S}\}}(Y^i(u)) Z_I(u) du \right)$$

$$606 \quad \tilde{N}_{I \rightarrow S} \left(\int_0^t \kappa_2 \tilde{X}_I^V(u) du \right) = \sum_{i=1}^V N_{\tilde{I}+I \rightarrow \tilde{S}+S}^i \left(\int_0^t \mathbb{1}_{\{\tilde{I}\}}(Y^i(u)) du \right).$$

607

608 Then,

$$609 \quad \left| \frac{X_S^V(t)}{V} - \frac{\tilde{X}_S^V(t)}{V} \right| \leq \Delta(t) + \frac{1}{V} \int_0^t \kappa_1 X_S^V(u) \left| \frac{X_I^V(u)}{V} - Z_I(u) \right| du$$

$$610 \quad + \int_0^t \kappa_1 \left| \frac{X_S^V(u)}{V} - \frac{\tilde{X}_S^V(u)}{V} \right| Z_I(u) du + \int_0^t \kappa_2 \left| \frac{X_I^V(u)}{V} - \frac{\tilde{X}_I^V(u)}{V} \right| du,$$

611

612

613 where

$$\begin{aligned}
614 \quad \Delta(t) &= \frac{1}{V} \left| N_{S+I \rightarrow 2I} \left(\int_0^t \frac{\kappa_1}{V} X_S^V(u) X_I^V(u) du \right) - \int_0^t \frac{\kappa_1}{V} X_S^V(u) X_I^V(u) du \right| \\
615 &\quad + \frac{1}{V} \left| N_{I \rightarrow S} \left(\int_0^t \kappa_2 X_I^V(u) du \right) - \int_0^t \kappa_2 X_I^V(u) du \right| \\
616 &\quad + \frac{1}{V} \left| \tilde{N}_{S+I \rightarrow 2I} \left(\int_0^t \kappa_1 \tilde{X}_S^V(u) Z_I(u) du \right) - \int_0^t \kappa_1 \tilde{X}_S^V(u) Z_I(u) du \right| \\
617 &\quad + \frac{1}{V} \left| \tilde{N}_{I \rightarrow S} \left(\int_0^t \kappa_2 \tilde{X}_I^V(u) du \right) - \int_0^t \kappa_2 \tilde{X}_I^V(u) du \right|.
\end{aligned}$$

619 Using $X_I^V(t) = V - X_S^V(t)$ and $Z_I(t) \leq 1$ for all $0 \leq t \leq T$ we obtain

$$\begin{aligned}
620 \quad \left| \frac{X_S^V(t)}{V} - \frac{\tilde{X}_S^V(t)}{V} \right| &\leq \Delta(t) + \int_0^t \kappa_1 \left| \frac{X_I^V(u)}{V} - Z_I(u) \right| du \\
621 &\quad + \int_0^t (\kappa_1 + \kappa_2) \left| \frac{X_S^V(u)}{V} - \frac{\tilde{X}_S^V(u)}{V} \right| du.
\end{aligned}$$

624 By taking the supremum on $0 \leq t \leq T$ on both sides and by applying the Gronwall inequality, we have

$$626 \quad \sup_{0 \leq t \leq T} \left| \frac{X_S^V(t)}{V} - \frac{\tilde{X}_S^V(t)}{V} \right| \leq \left(\sup_{0 \leq t \leq T} \Delta(t) + \kappa_1 T \sup_{0 \leq t \leq T} \left| \frac{X_I^V(u)}{V} - Z_I(u) \right| \right) e^{(\kappa_1 + \kappa_2)T}.$$

627 For notational convenience, let $\nu = \varepsilon e^{-(\kappa_1 + \kappa_2)T}$. Hence, (5.7) is smaller than

$$628 \quad (5.8) \quad P \left(\sup_{0 \leq t \leq T} \Delta(t) > \frac{\nu}{2} \right) + P \left(\sup_{0 \leq t \leq T} \left| \frac{X_I^V(u)}{V} - Z_I(u) \right| > \frac{\nu}{2\kappa_1 T} \right).$$

629 By noting that $P(\sup_{0 \leq t \leq T} \Delta(t) > \nu/2)$ is smaller than

$$\begin{aligned}
630 \quad &P \left(\sup_{0 \leq t \leq T} \frac{1}{V} \left| N_{S+I \rightarrow 2I} \left(\int_0^t \frac{\kappa_1}{V} X_S^V(u) X_I^V(u) du \right) - \int_0^t \frac{\kappa_1}{V} X_S^V(u) X_I^V(u) du \right| > \frac{\nu}{8} \right) \\
631 &+ P \left(\sup_{0 \leq t \leq T} \frac{1}{V} \left| N_{I \rightarrow S} \left(\int_0^t \kappa_2 X_I^V(u) du \right) - \int_0^t \kappa_2 X_I^V(u) du \right| > \frac{\nu}{8} \right) \\
632 &+ P \left(\sup_{0 \leq t \leq T} \frac{1}{V} \left| \tilde{N}_{S+I \rightarrow 2I} \left(\int_0^t \kappa_1 \tilde{X}_S^V(u) Z_I(u) du \right) - \int_0^t \kappa_1 \tilde{X}_S^V(u) Z_I(u) du \right| > \frac{\nu}{8} \right) \\
633 &+ P \left(\sup_{0 \leq t \leq T} \frac{1}{V} \left| \tilde{N}_{I \rightarrow S} \left(\int_0^t \kappa_2 \tilde{X}_I^V(u) du \right) - \int_0^t \kappa_2 \tilde{X}_I^V(u) du \right| > \frac{\nu}{8} \right),
\end{aligned}$$

635 we obtain that (5.8) is smaller than

$$\begin{aligned}
636 & \\
637 \quad &12 \exp \left(\frac{\kappa_1 \varepsilon T}{2} - \frac{\nu}{24} \sqrt{V} \right) + 12 \exp \left(\frac{\kappa_2 \varepsilon T}{2} - \frac{\nu}{24} \sqrt{V} \right) \\
638 &+ 6 \exp \left(\frac{\kappa_1 \varepsilon T}{2} \left(1 + \frac{\nu}{\kappa_1 T} \right)^2 + \frac{\kappa_2 \varepsilon T}{2} \left(1 + \frac{\nu}{\kappa_1 T} \right) - \frac{\nu}{12\kappa_1 T} e^{-T(\kappa_1 - \kappa_2) - \nu} \sqrt{V} \right) \\
639 &
\end{aligned}$$

640 by Lemma A.1 and Theorem A.2 (for the special case of the SIS model, see Exam-
 641 ple A.3). We note that $\exp(h)$ is defined as e^h for all real numbers h . It follows
 642 that (5.7) tends to 0 as V tends to ∞ with the same rate as $e^{-C\sqrt{V}}$ for some positive
 643 constant C . This is always the case, and bounds for more general models are provided
 644 by Theorem A.5.

645 Appendix A. Proofs and explicit bounds.

646 In this section we give proofs for the results stated above, together with more
 647 precise bounds on the quantities of interest. To this aim, we first define the following
 648 quantities: for all $V \in \mathbb{Z}_{\geq 1}$ and $\varepsilon \in \mathbb{R}_{>0}$ let

$$649 \mathcal{A}_{V,\varepsilon,t} = \left\{ \sup_{u \in [0,t]} \left\| \frac{X^V(u)}{V} - Z(u) \right\|_{\infty} \leq \varepsilon \right\} \quad \text{and} \quad p^{V,\varepsilon,t} = P(\mathcal{A}_{V,\varepsilon,t}^c) = 1 - P(\mathcal{A}_{V,\varepsilon,t}),$$

650 where the superscript “c” denotes the complement. Note that, for any fixed V and ε ,
 651 the sequence of events $\mathcal{A}_{V,\varepsilon,t}$ is monotone in t , and $p^{V,\varepsilon,t}$ is a non-decreasing
 652 function of t attaining its maximum for the value $t = T$.

653 Define the $\mathbb{Z}_{\geq 0}^d$ -valued process $X^{V,\varepsilon}$ on $[0, T]$ in the following way: for any $S \in \mathcal{X}$
 654 and any $t \in [0, T]$, let

$$655 (A.1) \quad X_S^{V,\varepsilon}(t) = \min\{\max\{X_S^V(t), VZ_S(t) - V\varepsilon\}, VZ_S(t) + V\varepsilon\}.$$

656 Hence, by definition for all $t \in \mathbb{R}_{>0}$

$$657 \left\| \frac{X^{V,\varepsilon}(t)}{V} - Z(t) \right\|_{\infty} \leq \varepsilon.$$

658 Moreover, define the process $\hat{X}^{V,\varepsilon}$ by

$$659 \hat{X}^{V,\varepsilon}(t) = X^V(0) + \sum_{y \rightarrow y' \in \mathcal{R}} (y' - y) N_{y \rightarrow y'} \left(\int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \right)$$

660 for all $t \in [0, T]$, where the processes $N_{y \rightarrow y'}$ are the same as in (4.6). Note that for
 661 any $t \in [0, T]$ we have $\mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}} X^{V,\varepsilon}(t) = \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}} X^V(t) = \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}} \hat{X}^{V,\varepsilon}(t)$. In particular,
 662 it follows that

$$663 \sup_{0 \leq u \leq t} \left\| \frac{X^{V,\varepsilon}(u)}{V} - Z(u) \right\|_{\infty} \leq \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}} \sup_{0 \leq u \leq t} \left\| \frac{\hat{X}^{V,\varepsilon}(u)}{V} - Z(u) \right\|_{\infty} + \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}^c} \varepsilon$$

$$664 (A.2) \quad \leq \sup_{0 \leq u \leq t} \left\| \frac{\hat{X}^{V,\varepsilon}(u)}{V} - Z(u) \right\|_{\infty}.$$

666 For any $t \in [0, T]$ and any $\varepsilon \in \mathbb{R}_{>0}$ let

$$667 \Omega_1^{\varepsilon,t} = \{Z(u) + h : u \in [0, t], h \in \mathbb{R}^d, \|h\|_{\infty} \leq \varepsilon\} \cap \mathbb{R}_{\geq 0}^d$$

668 be the (one-dimensional) neighbourhood of the solution Z on the interval $[0, t]$ with
 669 amplitude ε , intersected with the non-negative orthant. Note that for all $t \in [0, T]$ we
 670 have $X^{V,\varepsilon}(t)/V \in \Omega_1^{\varepsilon,V}$. Similarly, let

$$671 \Omega_2^{\varepsilon,t} = \{(Z(u) + h, Z(u) + h') : u \in [0, t], h, h' \in \mathbb{R}^d, \|h\|_{\infty} \leq \varepsilon, \|h'\|_{\infty} \leq \varepsilon\} \cap \mathbb{R}_{\geq 0}^{2d}$$

672 be the two-dimensional neighbourhood of the Z restricted to $[0, t]$ with amplitude ε ,
 673 intersected with the non-negative orthant.

674 To conclude, it is convenient to introduce in this section a notation for centered
 675 Poisson processes: given a Poisson process N , we denote by \bar{N} the process defined
 676 by $\bar{N}(t) = N(t) - t$ for all $t \in \mathbb{R}_{\geq 0}$. In order to bound $p^{V, \varepsilon, t}$ from above and prove
 677 Theorem 5.8 we need the following results concerning centered Poisson processes. For
 678 completeness, we provide a proof as we were not able to find it in the literature, even
 679 if small variations of Lemma A.1 are well-known and obtained as an application of
 680 Doob's inequality or Kolmogorov's maximal inequality.

681 LEMMA A.1. *Let N be a Poisson process and let $T, \varepsilon \in \mathbb{R}_{>0}$. Then, for all $n \in$
 682 $\mathbb{Z}_{\geq 1}$*

$$683 \quad P \left(\sup_{t \in [0, nT]} \left| \frac{\bar{N}(t)}{n} \right| > \varepsilon \right) \leq 6 \exp \left(\frac{e}{2} T - \frac{\varepsilon \sqrt{n}}{3} \right).$$

684 *Proof.* For all $j \in \mathbb{Z}_{\geq 1}$ and all $h \in \mathbb{R}_{>0}$ define

$$685 \quad (\text{A.3}) \quad \Xi_j^h = \bigcup_{i=0}^{2^j h} \left\{ \frac{i}{2^j} \right\}.$$

686 Since \bar{N} is almost surely right continuous, we have that for all $n \in \mathbb{Z}_{\geq 1}$ and all
 687 $T \in \mathbb{R}_{>0}$

$$688 \quad \sup_{t \in [0, nT]} \left| \frac{\bar{N}(t)}{n} \right| = \lim_{j \rightarrow \infty} \max_{t \in \Xi_j^{nT}} \left| \frac{\bar{N}(t)}{n} \right|$$

689 almost surely. Since for all $j \in \mathbb{Z}_{\geq 1}$ we have $\Xi_j^{nT} \subset \Xi_{j+1}^{nT}$, by continuity of the
 690 probability measure we have

$$691 \quad P \left(\sup_{t \in [0, nT]} \left| \frac{\bar{N}(t)}{n} \right| > \varepsilon \right) = \lim_{j \rightarrow \infty} P \left(\max_{t \in \Xi_j^{nT}} \left| \frac{\bar{N}(t)}{n} \right| > \varepsilon \right).$$

692 By Etemadi's inequality we have

$$693 \quad P \left(\max_{t \in \Xi_j^{nT}} \left| \frac{\bar{N}(t)}{n} \right| > \varepsilon \right) \leq 3 \max_{t \in \Xi_j^{nT}} P \left(\left| \frac{\bar{N}(t)}{n} \right| > \frac{\varepsilon}{3} \right).$$

694 Moreover, for any real $\beta \in (0, 1)$ and any real $t \in (0, nT)$ we have

$$\begin{aligned} 695 \quad P \left(\left| \frac{\bar{N}(t)}{n} \right| > \frac{\varepsilon}{3} \right) &\leq P \left(\frac{\bar{N}(t)}{n} > \frac{\varepsilon}{3} \right) + P \left(-\frac{\bar{N}(t)}{n} > \frac{\varepsilon}{3} \right) \\ 696 \quad &= P \left(e^{\frac{n^\beta \bar{N}(t)}{n}} > e^{\frac{n^\beta \varepsilon}{3}} \right) + P \left(e^{-\frac{n^\beta \bar{N}(t)}{n}} > e^{\frac{n^\beta \varepsilon}{3}} \right) \\ 697 \quad &\leq 2 \exp \left(-\frac{n^\beta \varepsilon}{3} \right) \exp \left(t(e^{n^{\beta-1}} - 1 - n^{\beta-1}) \right) \\ 698 \quad &\leq 2 \exp \left(-\frac{n^\beta \varepsilon}{3} \right) \exp \left(nT \frac{n^{2\beta-2}}{2} e^{n^{\beta-1}} \right), \\ 699 \quad &\leq 2 \exp \left(-\frac{n^\beta \varepsilon}{3} \right) \exp \left(nT \frac{n^{2\beta-2}}{2} e \right), \\ 700 \end{aligned}$$

701 where the inequality in the third line follows from the Markov's inequality and the
 702 known form of the moment generating function of a Poisson random variable, which
 703 leads to $E[e^{n^{\beta-1}\bar{N}(t)}] = e^{-n^{\beta-1}t}e^{t(e^{n^{\beta-1}}-1)}$ and $E[e^{-n^{\beta-1}\bar{N}(t)}] = e^{n^{\beta-1}t}e^{t(e^{-n^{\beta-1}}-1)}$.
 704 Hence, for all $n \in \mathbb{Z}_{\geq 1}$ we have that both $E[e^{n^{\beta-1}\bar{N}(t)}]$ and $E[e^{-n^{\beta-1}\bar{N}(t)}]$ are less than
 705 or equal to $e^{t(e^{n^{\beta-1}}-1-n^{\beta-1})}$. The inequality in the fourth line derives from the Taylor
 706 expansion of the exponential function. By choosing $\beta = 1/2$ we have

$$707 \quad P\left(\left|\frac{\bar{N}(t)}{n}\right| > \frac{\varepsilon}{3}\right) \leq 2 \exp\left(-\frac{\varepsilon\sqrt{n}}{3}\right) \exp\left(\frac{e}{2}T\right),$$

708 which completes the proof. \square

709 **A.1. Estimates for $p^{V,\varepsilon,t}$.** Many papers have focused on quantifying the dis-
 710 tance between the process X^V and its fluid limit Z . Among these, we list [1–3, 13,
 711 17, 20] with no claim of completeness. Here we use Lemma A.1 to show the following
 712 upper bound on $p^{V,\varepsilon,t}$. While similar estimates are known in the reaction network
 713 community, we give a formal proof of the bound we propose as we could not find it
 714 in the literature. Before stating the result, we define the following quantities:

$$715 \quad R = \max_{y \rightarrow y' \in \mathcal{R}} \|y' - y\|_{\infty},$$

$$716 \quad \Lambda_0^{\varepsilon,t} = \sup_{z \in \Omega_1^{\varepsilon,t}} \sum_{y \rightarrow y' \in \mathcal{R}} \lambda_{y \rightarrow y'}(z), \quad \Lambda_1^{\varepsilon,t} = \int_0^t \Lambda_0^{\varepsilon,u} du$$

$$717 \quad L_0^{\varepsilon,t} = \sup_{\substack{(z,z') \in \Omega_2^{\varepsilon,t} \\ z \neq z'}} \sum_{y \rightarrow y' \in \mathcal{R}} \frac{|\lambda_{y \rightarrow y'}(z) - \lambda_{y \rightarrow y'}(z')|}{\|z - z'\|_{\infty}}, \quad L_1^{\varepsilon,t} = \int_0^t L_0^{\varepsilon,u} du$$

$$718 \quad \delta_0^{V,\varepsilon,t} = \sup_{z \in \Omega_1^{\varepsilon,t}} \sum_{y \rightarrow y' \in \mathcal{R}} \left| \frac{\lambda_{y \rightarrow y'}^V(\lfloor Vz \rfloor)}{V} - \lambda_{y \rightarrow y'}(z) \right|, \quad \delta_1^{V,\varepsilon,t} = \int_0^t \delta_0^{V,\varepsilon,u} du$$

$$719 \quad \eta^{V,\varepsilon,t}(\gamma) = e^{-L_1^{2\varepsilon,t}} \gamma \varepsilon - \delta_1^{V,2\varepsilon,t},$$

721 where in the last definition γ is any real number in $(0, 1]$. Note that $\Lambda_0^{\varepsilon,t}$ and $\delta_0^{V,\varepsilon,t}$ are
 722 finite for any $t \in [0, T]$, since the solution Z exists up to time T and the functions $\lambda_{y \rightarrow y'}$
 723 are locally Lipschitz by Assumption 3.1. The local Lipschitzianity of the functions
 724 $\lambda_{y \rightarrow y'}$ also implies that $L_0^{\varepsilon,t}$ is finite for all $\varepsilon \in \mathbb{R}_{>0}$ and $t \in [0, T]$. It also follows
 725 from Assumption 3.1 that $\delta_0^{V,\varepsilon,t}$ tends to zero as V tends to infinity. Furthermore,
 726 note that for fixed $V \in \mathbb{Z}_{\geq 1}$ and $\varepsilon \in \mathbb{R}_{>0}$, the quantities $\Lambda_0^{\varepsilon,t}$, $L_0^{\varepsilon,t}$, and $\delta_0^{V,\varepsilon,t}$ are
 727 all non-decreasing functions of t . As a consequence, for all $t \in [0, T]$, $\varepsilon \in \mathbb{R}_{>0}$, and
 728 $V \in \mathbb{Z}_{\geq 1}$ we have

$$729 \quad \Lambda_1^{\varepsilon,t} \leq t\Lambda_0^{\varepsilon,t}, \quad L_1^{\varepsilon,t} \leq tL_0^{\varepsilon,t}, \quad \text{and} \quad \delta_1^{V,\varepsilon,t} \leq t\delta_0^{V,\varepsilon,t}.$$

730 It follows that for all $t \in [0, T]$, $\varepsilon \in \mathbb{R}_{>0}$, and $\gamma \in (0, 1]$ the quantity $\eta^{V,\varepsilon,t}(\gamma)$ tends to
 731 the positive quantity $e^{-L_1^{2\varepsilon,t}} \gamma \varepsilon$ as V tends to infinity. We can now state the following
 732 theorem.

733 **THEOREM A.2.** *For any $\varepsilon, t \in \mathbb{R}_{>0}$, any $\gamma \in (0, 1]$, and any $V \in \mathbb{Z}_{\geq 1}$ large*
 734 *enough such that $\eta^{V,2\varepsilon,t}(\gamma) > 0$, we have*

$$735 \quad p^{V,\varepsilon,t} \leq p^{V,(1-\gamma)\varepsilon e^{-L_1^{2\varepsilon,t}},0} + 6 \exp\left(\frac{e}{2}\Lambda_1^{2\varepsilon,t} + \frac{e}{2}\delta_1^{V,2\varepsilon,t} - \frac{1}{3R}\eta^{V,\varepsilon,t}(\gamma)\sqrt{V}\right)$$

736 *Proof.* First, note that

$$737 \quad p^{V,\varepsilon,t} = P \left(\sup_{u \in [0,t]} \left\| \frac{X^V(u)}{V} - Z(u) \right\|_\infty > \varepsilon \right) = P \left(\sup_{u \in [0,t]} \left\| \frac{X^{V,2\varepsilon}(u)}{V} - Z(u) \right\|_\infty > \varepsilon \right)$$

$$738 \quad = P \left(\sup_{u \in [0,t]} \left\| \frac{\hat{X}^{V,2\varepsilon}(u)}{V} - Z(u) \right\|_\infty > \varepsilon \right).$$

740 Moreover, by superposition, for all $V \in \mathbb{Z}_{\geq 1}$ and all $\varepsilon \in \mathbb{R}_{>0}$ we can define a unit-rate
741 Poisson process $U^{V,2\varepsilon}$ coupled with X^V in such a way that for all $t \in \mathbb{R}_{\geq 0}$

$$742 \quad U^{V,2\varepsilon} \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^t \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(u)) du \right) = \sum_{y \rightarrow y' \in \mathcal{R}} N_{y \rightarrow y'} \left(\int_0^t \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(u)) du \right).$$

743 Hence, by using (2.2) we have

$$744 \quad \left\| \frac{\hat{X}^{V,2\varepsilon}(u)}{V} - Z(u) \right\|_\infty \leq \left\| \frac{\hat{X}^{V,2\varepsilon}(0)}{V} - Z(0) \right\|_\infty + \frac{R}{V} \left| \sum_{y \rightarrow y' \in \mathcal{R}} \bar{N}_{y \rightarrow y'} \left(\int_0^u \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(w)) dw \right) \right|$$

$$745 \quad + \int_0^u \left| \sum_{y \rightarrow y' \in \mathcal{R}} \left(\frac{\lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(w))}{V} - \lambda_{y \rightarrow y'} \left(\frac{X^{V,2\varepsilon}(w)}{V} \right) \right) dw \right|$$

$$746 \quad + \int_0^u \left| \sum_{y \rightarrow y'} \left(\lambda_{y \rightarrow y'} \left(\frac{X^{V,2\varepsilon}(w)}{V} \right) - \lambda_{y \rightarrow y'}(Z(w)) \right) dw \right|$$

$$747 \quad \leq \left\| \frac{X^V(0)}{V} - Z(0) \right\|_\infty + \frac{R}{V} \left| \bar{U}^{V,2\varepsilon} \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^u \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(w)) dw \right) \right|$$

$$748 \quad + \delta_1^{V,2\varepsilon,u} + \int_0^u L_0^{2\varepsilon,w} \left\| \frac{X^{V,2\varepsilon}(w)}{V} - Z(w) \right\|_\infty dw$$

750 By using (A.2), by taking the supremum over $[0, t]$ on both sides we obtain

$$751 \quad \sup_{0 \leq u \leq t} \left\| \frac{\hat{X}^{V,2\varepsilon}(u)}{V} - Z(u) \right\|_\infty \leq \left\| \frac{X^V(0)}{V} - Z(0) \right\|_\infty$$

$$752 \quad + \frac{R}{V} \sup_{0 \leq u \leq t} \left| \bar{U}^{V,2\varepsilon} \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^u \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(w)) dw \right) \right|$$

$$753 \quad + \delta_1^{V,2\varepsilon,t} + \int_0^t L_0^{2\varepsilon,u} \sup_{0 \leq w \leq u} \left\| \frac{\hat{X}^{V,2\varepsilon}(w)}{V} - Z(w) \right\|_\infty du.$$

754
755 By Gronwall's inequality we get

$$756 \quad \sup_{0 \leq u \leq t} \left\| \frac{\hat{X}^{V,2\varepsilon}(t)}{V} - Z(t) \right\|_\infty \leq e^{L_1^{2\varepsilon,t}} \left\| \frac{X^V(0)}{V} - Z(0) \right\|_\infty$$

$$757 \quad + \frac{Re^{L_1^{2\varepsilon,t}}}{V} \sup_{0 \leq u \leq t} \left| \bar{U}^{V,2\varepsilon} \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^u \lambda_{y \rightarrow y'}^V(X^{V,2\varepsilon}(w)) dw \right) \right|$$

$$758 \quad + e^{L_1^{2\varepsilon,t}} \delta_1^{V,2\varepsilon,t}.$$

760 By noting that for all $t \in \mathbb{R}_{\geq 0}$

$$761 \quad \sup_{z \in \Omega_1^{2\varepsilon, t}} \sum_{y \rightarrow y' \in \mathcal{R}} \frac{\lambda_{y \rightarrow y'}^V([Vz])}{V} \leq \Lambda_0^{2\varepsilon, t} + \delta_0^{V, 2\varepsilon, t},$$

762 we get

$$763 \quad p^{V, \varepsilon, t} \leq P \left(e^{L_1^{2\varepsilon, t}} \left\| \frac{X^V(0)}{V} - Z(0) \right\|_{\infty} > (1 - \gamma)\varepsilon \right) \\ 764 \quad + P \left(\operatorname{Re}^{L_1^{2\varepsilon, t}} \sup_{0 \leq u \leq V(\Lambda_1^{2\varepsilon, t} + \delta_1^{V, 2\varepsilon, t})} \left| \frac{\bar{U}^{V, 2\varepsilon}(u)}{V} \right| + e^{L_1^{2\varepsilon, t}} \delta_1^{V, 2\varepsilon, t} > \gamma\varepsilon \right)$$

766 for any γ in $(0, 1]$. the proof is concluded by Lemma A.1. \square

767 *Example A.3.* Consider the SIS reaction network described in (4.2). In this case,
768 in accordance with the classical mass-action choice of kinetics we have

$$769 \quad \lambda_{S+I \rightarrow 2I}^V(x) = \frac{1}{V} \kappa_1 x_S x_I \quad \text{and} \quad \lambda_{I \rightarrow S}^V(x) = \kappa_2 x_I$$

770 for some positive constants κ_1 and κ_2 . Hence, Assumption 3.1 is satisfied with

$$771 \quad \lambda_{S+I \rightarrow 2I}(z) = \kappa_1 z_S z_I \quad \text{and} \quad \lambda_{I \rightarrow S}(z) = \kappa_2 z_I.$$

772 The corresponding solution Z exists for all non-negative times t , for all initial condi-
773 tions $Z(0) = z^*$. Moreover, note that the sum of infected and susceptible individuals
774 is kept constant, hence for all $t \in \mathbb{R}_{> 0}$ we have $Z_S(t) + Z_I(t) = z_S^* + z_I^* = \|z^*\|_1$. In
775 this case we can obtain the following rough estimates

$$776 \quad R = 2, \quad \Lambda_0^{\varepsilon, t} \leq (\|z^*\|_1 + \varepsilon)[\kappa_1(\|z^*\|_1 + \varepsilon) + \kappa_2], \quad L_0^{\varepsilon, t} \leq \kappa_1(\|z^*\|_1 + \varepsilon) + \kappa_2, \\ 777 \quad \delta_0^{V, \varepsilon, t} = 0, \quad \eta^{V, \varepsilon, t} \geq \varepsilon e^{-t\kappa_1(\|z^*\|_1 + 2\varepsilon) + t\kappa_2}.$$

779 If we assume $X^V(0) = Vz^*$, then $p^{V, 0, 0} = 0$. It follows from Theorem A.2 with the
780 choice $\gamma = 1$ that in this case

$$781 \quad p^{V, \varepsilon, t} \leq 6 \exp \left(\frac{t}{2} (\|z^*\|_1 + 2\varepsilon) [\kappa_1(\|z^*\|_1 + 2\varepsilon) + \kappa_2] - \frac{\varepsilon \sqrt{V}}{6} e^{-t[\kappa_1(\|z^*\|_1 + 2\varepsilon) - \kappa_2]} \right),$$

782 where $\exp(h)$ is defined as e^h for all real numbers h .

783 **A.2. Proof of Theorem 5.2.** First of all, we define some quantities that are
784 useful to give specific bounds on our approximation error. Define

$$785 \quad \tilde{\Lambda}_0^t = \max_{\tilde{S} \in \tilde{\mathcal{X}}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, Z(t)), \\ 786 \quad \tilde{L}_0^{\varepsilon, t} = \sup_{\substack{(z, z') \in \Omega_2^{\varepsilon, t} \\ z \neq z'}} \max_{\tilde{S} \in \tilde{\mathcal{X}}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \frac{|\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z')|}{\|z - z'\|_{\infty}} \\ 787 \quad \tilde{\delta}_0^{\varepsilon, t} = \sup_{z \in \Omega_1^{\varepsilon, t}} \max_{\tilde{S} \in \tilde{\mathcal{X}}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} |\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\tilde{S}, [Vz]) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z)| \\ 788 \quad \tilde{\Lambda}_1^t = \int_0^t \tilde{\Lambda}_0^u du, \quad \tilde{L}_1^{\varepsilon, t} = \int_0^t L_0^{\varepsilon, u} du, \quad \tilde{\delta}_1^{\varepsilon, t} = \int_0^t \delta_0^{V, \varepsilon, u} du. \\ 789$$

790 Note that $\tilde{\Lambda}_0^t$ is finite for any $t \in [0, T]$, due to the fact that Z is defined over the
 791 whole interval $[0, T]$. Moreover the functions $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}$ are locally Lipschitz on $\mathbb{R}_{>0}^d$
 792 by Lemma 5.1, hence $\tilde{L}_0^{\varepsilon, t}$ is finite for all $t \in [0, T]$. Finally, $\tilde{\delta}_0^{V, \varepsilon, t}$ is finite for all
 793 $t \in [0, T]$ by Lemma 5.1. Note that, for fixed V and ε , the quantities $\tilde{L}_0^{\varepsilon, t}$ and $\tilde{\delta}_0^{\varepsilon, t}$
 794 are non-decreasing functions of t . As a consequence, for all $t \in [0, T]$, $\varepsilon \in \mathbb{R}_{>0}$, and
 795 $V \in \mathbb{Z}_{\geq 1}$ we have

$$796 \quad (\text{A.4}) \quad \tilde{\Lambda}_1^t \leq t\tilde{\Lambda}_0^t, \quad \tilde{L}_1^{\varepsilon, t} \leq t\tilde{L}_0^{\varepsilon, t}, \quad \text{and} \quad \tilde{\delta}_1^{\varepsilon, t} \leq t\tilde{\delta}_0^{\varepsilon, t}.$$

797 Before proving Theorem 5.2 we show the following stronger result.

798 **THEOREM A.4.** *Assume that Assumption 3.1 holds. Furthermore, assume that*
 799 *the random variables $X^V(0)/V$ converge in probability to a constant z^* as V goes to*
 800 *infinity. Assume that the solution Z to (2.2) with $Z(0) = z^*$ exists over the interval*
 801 *$[0, T]$ and that*

$$802 \quad m = \min_{\substack{S \in \mathcal{X} \\ u \in [0, T]}} Z_S(u) > 0.$$

803 Finally, assume that $Y^V(0) = Y(0)$ for all positive integers V . Then,

$$804 \quad (\text{A.5}) \quad P(Y^V(t) \neq Y(t)) = E[\|Y^V(t) - Y(t)\|_\infty].$$

805 Moreover, for any $0 < \varepsilon < m$

$$806 \quad \sup_{t \in [0, T]} E[\|Y^V(t) - Y(t)\|_\infty] \leq p^{V, \varepsilon, T} + (\tilde{\delta}_1^{V, \varepsilon, T} + \varepsilon \tilde{L}_1^{\varepsilon, T}) e^{2\tilde{\Lambda}_1^T}.$$

807 *Proof.* First, note that

$$808 \quad (\text{A.6}) \quad \|Y^V(t) - Y(t)\|_\infty = \begin{cases} 1 & \text{if } Y^V(t) \neq Y(t) \\ 0 & \text{if } Y^V(t) = Y(t) \end{cases},$$

809 hence (A.5) holds. Consider the process

$$810 \quad \hat{Y}^V(t) = Y(0) + \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} (\tilde{S}' - \tilde{S}) N_{\tilde{S}+y \rightarrow \tilde{S}'+y'} \left(\int_0^t \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\hat{Y}^V(u), X^{V, \varepsilon}(u)) du \right). \quad \blacksquare$$

811 By equations (5.2) and (A.7), using the triangular inequality, we obtain

$$812 \quad E[\|\hat{Y}^V(t) - Y(t)\|_\infty] \\ 813 \quad \leq E \left[\int_0^t \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\hat{Y}^V(u), X^{V, \varepsilon}(u)) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(Y(u), Z(u)) \right| du \right] \\ 814 \quad \leq \Upsilon_1 + \Upsilon_2 + \Upsilon_3 \quad \blacksquare$$

816 where

$$\begin{aligned}
817 \quad \Upsilon_1 &= E \left[\int_0^t \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\hat{Y}^V(u), X^{V,\varepsilon}(u)) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'} \left(\hat{Y}^V(u), \frac{X^{V,\varepsilon}(u)}{V} \right) \right| du \right] \\
818 \quad \Upsilon_2 &= E \left[\int_0^t \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'} \left(\hat{Y}^V(u), \frac{X^{V,\varepsilon}(u)}{V} \right) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\hat{Y}^V(u), Z(u)) \right| du \right] \\
819 \quad \Upsilon_3 &= E \left[\int_0^t \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\hat{Y}^V(u), Z(u)) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(Y(u), Z(u)) \right| du \right]
\end{aligned}$$

821 Since for every $\tilde{S} + y \rightarrow \tilde{S}' + y' \in \tilde{\mathcal{R}}$ we have

$$\begin{aligned}
822 \quad \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(w, x) &= \mathbb{1}_{\{\tilde{S}\}}(w) \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\tilde{S}, x) \quad \text{for all } x \in \mathbb{Z}_{\geq 0}^d, w \in \tilde{\mathcal{X}} \\
823 \quad \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(w, z) &= \mathbb{1}_{\{\tilde{S}\}}(w) \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z) \quad \text{for all } z \in \mathbb{R}_{\geq 0}^d, w \in \tilde{\mathcal{X}},
\end{aligned}$$

825 we can write $\Upsilon_1 \leq \tilde{\delta}_1^{V,\varepsilon,t}$. Similarly, $\Upsilon_2 \leq \varepsilon \tilde{L}_1^{\varepsilon,t}$. Finally,

$$\begin{aligned}
826 \quad \Upsilon_3 &= E \left[\int_0^t \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \mathbb{1}_{\{\tilde{S}\}}(\hat{Y}^V(u)) - \mathbb{1}_{\{\tilde{S}\}}(Y(u)) \right| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, Z(u)) du \right] \\
827 \quad &\leq E \left[\int_0^t \sum_{\tilde{S} \in \tilde{\mathcal{X}}} \left| \mathbb{1}_{\{\tilde{S}\}}(\hat{Y}^V(u)) - \mathbb{1}_{\{\tilde{S}\}}(Y(u)) \right| \tilde{\Lambda}_0^u du \right] \\
828 \quad &= \int_0^t 2P(Y^V(u) \neq Y(u)) \tilde{\Lambda}_0^u du = 2 \int_0^t E \left[\|\hat{Y}^V(u) - Y(u)\|_\infty \right] \tilde{\Lambda}_0^u du,
\end{aligned}$$

830 where in the last equality we used (A.5). In conclusion,

$$831 \quad E \left[\|\hat{Y}^V(t) - Y(t)\|_\infty \right] \leq (\tilde{\delta}_1^{V,\varepsilon,t} + \varepsilon \tilde{L}_1^{\varepsilon,t}) + 2 \int_0^t E \left[\|\hat{Y}^V(u) - Y(u)\|_\infty \right] \tilde{\Lambda}_0^u du.$$

832 By the Gronwall inequality we then have

$$833 \quad E \left[\|\hat{Y}^V(t) - Y(t)\|_\infty \right] \leq (\tilde{\delta}_1^{V,\varepsilon,t} + \varepsilon \tilde{L}_1^{\varepsilon,t}) e^{2\tilde{\Lambda}_1^t}.$$

834 The result follows by taking the sup over $t \in [0, T]$ on both sides (the quantity
835 on the right-hand side of the inequality is non-decreasing in t) and by noting that
836 $\mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}} \hat{Y}^V(t) = \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}} Y^V(t)$ for all $t \in [0, T]$. Hence,

$$\begin{aligned}
837 \quad \|\hat{Y}^V(t) - Y(t)\|_\infty &= \|Y^V(t) - Y(t)\|_\infty \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}^c} + \|\hat{Y}^V(t) - Y(t)\|_\infty \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}} \\
838 \quad &\leq \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}^c} + \|\hat{Y}^V(t) - Y(t)\|_\infty \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}} \\
839 \quad &\leq \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}^c} + \|\hat{Y}^V(t) - Y(t)\|_\infty. \quad \square
\end{aligned}$$

841 We are now ready to prove Theorem 5.2

842 *Proof of Theorem 5.2.* It follows from Theorem A.4 that $P(Y^V(t) \neq Y(t)) =$
 843 $E[\|Y^V(t) - Y(t)\|_\infty]$. Moreover, for any $\varepsilon > 0$ we have $\lim_{V \rightarrow \infty} p^{V,\varepsilon,T} = 0$ by
 844 Theorem 3.2, and $\lim_{V \rightarrow \infty} \tilde{\delta}_1^{V,\varepsilon,T} = 0$ by Lemma 5.1 and (A.4). Hence,

$$845 \quad \lim_{V \rightarrow \infty} \sup_{t \in [0,T]} E[\|Y^V(t) - Y(t)\|_\infty] \leq \varepsilon \tilde{L}_1^{\varepsilon,T} e^{2\tilde{\Lambda}_1^T},$$

846 which concludes the proof by the arbitrariness of $\varepsilon > 0$ and by the fact that $\tilde{L}_0^{\varepsilon,T}$
 847 (hence $\tilde{L}_1^{\varepsilon,T}$) is non-decreasing in ε . \square

848 **A.3. Proof of Theorem 5.8.** Similarly to what was done in the previous section,
 849 we define the following quantities to give an upper bound for our approximation
 850 error. Define

$$851 \quad \hat{R} = \max_{y \rightarrow y' \in \mathcal{R}} \|\pi(y' - y)\|_\infty, \quad \hat{r} = \max_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left\| \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} - \frac{\tau(\tilde{S})}{\alpha(\tau(\tilde{S}))} \right\|_\infty,$$

$$852 \quad \hat{\Lambda}_0^t = \hat{r} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, Z(t)), \quad \hat{\Lambda}_1^t = \int_0^t \hat{\Lambda}_0^u du,$$

$$853 \quad \hat{\Lambda}_2^t = \max_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \int_0^t \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, Z(u)) du,$$

$$854 \quad \hat{\Lambda}_3^{V,\varepsilon,t} = \int_0^t \sup_{z \in \Omega_1^{\varepsilon,u}} \sum_{y \rightarrow y' \in \mathcal{R}} \frac{\lambda_{y \rightarrow y'}(\lfloor Vz \rfloor)}{V} du,$$

$$855 \quad \omega^{\varepsilon,t} = \hat{r} \sup_{\substack{(z,z') \in \Omega_2^{\varepsilon,t} \\ \|z-z'\|_\infty \leq \varepsilon}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \left| \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z) - \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}(\tilde{S}, z') \right|,$$

$$856 \quad \zeta^{\varepsilon,t} = \int_0^t (\|Z(u)\|_\infty + \varepsilon) du.$$

858 Note that $\hat{\Lambda}_0^t$, $\hat{\Lambda}_2^t$, and $\zeta^{\varepsilon,t}$ are finite for any $t \in [0, T]$, because Z is defined
 859 over the whole interval $[0, T]$ and the functions $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}$ are continuous on $\mathbb{R}_{>0}^d$ by
 860 Lemma 5.1. Lemma 5.1 also implies that $\omega^{\varepsilon,t}$ is finite for all $t \in [0, T]$ and $\varepsilon \in \mathbb{R}_{>0}$.
 861 Finally, $\hat{\Lambda}_3^{V,\varepsilon,t}$ is finite by Assumption 3.1. Note that, for fixed V and ε , the quantities
 862 $\hat{\Lambda}_3^{V,\varepsilon,t}$, $\omega^{\varepsilon,t}$, and $\zeta^{\varepsilon,t}$ are non-decreasing functions of t .

863 We now state and prove the following result, which immediately implies Theorem
 864 5.8. Note that $\delta_1^{V,\varepsilon,t}$ is as defined in Section A.1.

865 **THEOREM A.5.** *Consider a family of stochastic reaction systems with trackable*
 866 *species (Y^V, X^V) , and assume that Assumptions 3.1 and 5.6 are satisfied. Let $z^* \in$*
 867 *$\mathbb{R}_{>0}^d$ and $\tilde{X}^V(0) = \lfloor Vz^* \rfloor$. Define the process \tilde{X}^V by*

$$868 \quad \tilde{X}^V(t) = \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \frac{\tau(Y^{\tilde{S},i}(t))}{\alpha(\tau(Y^{\tilde{S},i}(t)))},$$

869 where the processes $(Y^{\tilde{S},i})_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}, i \in \mathbb{Z}_{\geq 1}}$ are independent and satisfy

$$870 \quad Y^{\tilde{S},i}(t) = \tilde{S} + \sum_{\tilde{S}'+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} (\tilde{S}'' - \tilde{S}') N_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}^{\tilde{S},i} \left(\int_0^t \lambda_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}(Y(u)^{\tilde{S},i}, Z(u)) du \right),$$

871 for a family of independent, identically distributed unit-rate Poisson processes $\{N_r^{\tilde{S},i}\}_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}, i \in \mathbb{Z}_{\geq 1}, r \in \tilde{\mathcal{R}}}$ ■
 872 For arbitrary $\nu_1, \nu_2, \nu_3 \in \mathbb{R}_{>0}$ define

$$873 \quad \nu = e^{\hat{\Lambda}_1^T} \left(\hat{R}\nu_1 + \hat{r}\nu_2 + \nu_3 + \hat{R}\delta_1^{V,\varepsilon,T} + \omega^{\varepsilon,T}\zeta^{\varepsilon,T} \right)$$

874 Then,

$$875 \quad P \left(\sup_{0 \leq t \leq T} \left\| \frac{\pi(X^V(t))}{V} - \frac{\tilde{X}^V(t)}{V} \right\|_{\infty} > \nu \right) \leq 6 \exp \left(\frac{e\hat{\Lambda}_3^{V,\varepsilon,t}}{2} - \frac{\nu_1\sqrt{V}}{3} \right)$$

$$877 \quad + 6 \exp \left(\frac{ec\hat{\Lambda}_2^t}{2} - \frac{\nu_2\sqrt{V}}{3} \right) + P \left(\left\| \frac{\pi(X^V(0))}{V} - \frac{\tilde{X}^V(0)}{V} \right\|_{\infty} > \nu_3 \right) + p^{V,\varepsilon,T},$$

878

879 where $c = \sum_{S \in \mathcal{X}} \alpha(S)z_S^*$.

880 *Proof.* By the superposition property of Poisson processes, for all $V \in \mathbb{Z}_{\geq 1}$ there
 881 exist two unit-rate Poisson processes U_1^V and U_2^V such that for all $t \in \mathbb{R}_{\geq 0}$

$$882 \quad U_1^V \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \right) = \sum_{y \rightarrow y' \in \mathcal{R}} N_{y \rightarrow y'} \left(\int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \right)$$

883 and

$$884 \quad U_2^V \left(\sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \int_0^t \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(u), Z(u)) du \right)$$

$$886 \quad = \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} N_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}^{\tilde{S},i} \left(\int_0^t \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(u), Z(u)) du \right)$$

887 ■

888 Note that

$$889 \quad \tilde{X}^V(t) = \tilde{X}^V(0) + \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \times$$

$$891 \quad \times N_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}^{\tilde{S},i} \left(\int_0^t \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y(u)^{\tilde{S},i}, Z(u)) du \right).$$

892

893 Hence, by triangular inequality,

$$894 \quad \sup_{0 \leq u \leq t} \left\| \frac{\pi(\hat{X}^{V,\varepsilon}(u))}{V} - \frac{\tilde{X}^V(u)}{V} \right\|_{\infty} \leq \left\| \frac{\pi(X^V(0))}{V} - \frac{\tilde{X}^V(0)}{V} \right\|_{\infty} + \sum_{i=1}^5 \Upsilon_i$$

895

896 where

$$\begin{aligned}
897 \quad \Upsilon_1 &= \sup_{0 \leq u \leq t} \sum_{y \rightarrow y' \in \mathcal{R}} \|\pi(y' - y)\|_\infty \frac{1}{V} \left| \bar{N}_{y \rightarrow y'} \left(\int_0^u \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(w)) dw \right) \right| \\
898 \quad &\leq \frac{\hat{R}}{V} \sup_{0 \leq u \leq t} \left| \bar{U}_1^V \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^u \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(w)) dw \right) \right| \\
899 \quad \Upsilon_2 &= \sup_{0 \leq u \leq t} \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \left\| \frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right\|_\infty \times \\
900 \quad &\times \frac{1}{V} \left| \bar{N}_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}^{\tilde{S},i} \left(\int_0^u \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(w), Z(w)) dw \right) \right| \\
901 \quad &\leq \frac{\hat{r}}{V} \sup_{0 \leq u \leq t} \left| \bar{U}_2^V \left(\sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \int_0^u \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(w), Z(w)) dw \right) \right| \\
902 \quad \Upsilon_3 &= \sup_{0 \leq u \leq t} \sum_{y \rightarrow y' \in \mathcal{R}} \|\pi(y' - y)\|_\infty \int_0^u \left| \frac{\lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(w))}{V} - \lambda_{y \rightarrow y'} \left(\frac{X^{V,\varepsilon}(w)}{V} \right) \right| dw \\
903 \quad &\leq \hat{R} \delta_1^{V,\varepsilon,t} \\
904 \quad \Upsilon_4 &= \sup_{0 \leq u \leq t} \left\| \sum_{y \rightarrow y' \in \mathcal{R}} \pi(y' - y) \int_0^u \lambda_{y \rightarrow y'} \left(\frac{X^{V,\varepsilon}(w)}{V} \right) dw \right. \\
905 \quad &\quad \left. - \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \int_0^u \frac{X_{\tau(\tilde{S}')}^{V,\varepsilon}(w)}{V} \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(\tilde{S}', Z(w)) dw \right\|_\infty \\
906 \quad \Upsilon_5 &= \sup_{0 \leq u \leq t} \left\| \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \int_0^u \frac{X_{\tau(\tilde{S}')}^{V,\varepsilon}(w)}{V} \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(\tilde{S}', Z(w)) dw \right. \\
907 \quad &\quad \left. - \frac{1}{V} \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \int_0^u \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(w), Z(w)) dw \right\|_\infty \blacksquare
\end{aligned}$$

909 We first focus on rewriting Υ_4 and Υ_5 . To this aim, first note that by identifying
910 species with canonical vectors of \mathbb{R}^d as previously done in the paper, we have that for
911 all $y \in \mathcal{C}$

$$912 \quad \pi(y) = \sum_{S \in \bar{\mathcal{X}}} y_S S = \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{y_{\tau(\tilde{S})} \tau(\tilde{S})}{\alpha(\tau(\tilde{S}))}.$$

913 Hence, for all $y \rightarrow y' \in \mathcal{R}$

$$\begin{aligned}
914 \quad \pi(y' - y) &= \sum_{\tilde{S}' \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{y_{\tau(\tilde{S}')\tau(\tilde{S}')}}{\alpha(\tau(\tilde{S}'))} - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{y_{\tau(\tilde{S})\tau(\tilde{S})}}{\alpha(\tau(\tilde{S}))} \\
915 &= \sum_{\tilde{S}' \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})p_{y \rightarrow y'}(\tilde{S}, \tilde{S}')} - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{y_{\tau(\tilde{S})}}{\alpha(\tau(\tilde{S}))} \tau(\tilde{S}), \\
916
\end{aligned}$$

917 where we used Assumption 5.6 in the last equality. By recalling that $\tau(\Delta) = 0$ and
918 $\sum_{\tilde{S}' \in \tilde{\mathcal{X}}} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}')$ for all $y \rightarrow y' \in \mathcal{R}$ and $\tilde{S} \in \tilde{\mathcal{X}}$, we further obtain

$$\begin{aligned}
919 \quad \pi(y' - y) &= \sum_{\tilde{S}' \in \tilde{\mathcal{X}}} \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} y_{\tau(\tilde{S})p_{y \rightarrow y'}(\tilde{S}, \tilde{S}')} \\
920 &\quad - \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \frac{y_{\tau(\tilde{S})}}{\alpha(\tau(\tilde{S}))} \tau(\tilde{S}) \sum_{\tilde{S}' \in \tilde{\mathcal{X}}} p_{y \rightarrow y'}(\tilde{S}, \tilde{S}') \\
921 &= \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}' \in \tilde{\mathcal{X}}} \left(\frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} - \frac{\tau(\tilde{S})}{\alpha(\tau(\tilde{S}))} \right) y_{\tau(\tilde{S})p_{y \rightarrow y'}(\tilde{S}, \tilde{S}')}. \\
922
\end{aligned}$$

923 It follows that

$$\begin{aligned}
924 \quad \sum_{y \rightarrow y' \in \mathcal{R}} \pi(y' - y) \int_0^u \lambda_{y \rightarrow y'} \left(\frac{X^{V,\varepsilon}(w)}{V} \right) dw \\
925 &= \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \int_0^u y_{\tau(\tilde{S}')p_{y \rightarrow y'}(\tilde{S}', \tilde{S}'')} \lambda_{y \rightarrow y'} \left(\frac{X^{V,\varepsilon}(w)}{V} \right) dw \\
926 &= \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \left(\frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right) \int_0^u \frac{X^{V,\varepsilon}_{\tau(\tilde{S}')} (w)}{V} \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'} \left(\tilde{S}', \frac{X^{V,\varepsilon}(w)}{V} \right) dw, \\
927
\end{aligned}$$

928 which in turn implies

$$\begin{aligned}
929 \quad \Upsilon_4 &\leq \sup_{0 \leq u \leq t} \sum_{\tilde{S}' + y \rightarrow \tilde{S}'' + y' \in \tilde{\mathcal{R}}} \left\| \frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right\|_{\infty} \times \\
930 &\quad \times \int_0^u \frac{X^{V,\varepsilon}_{\tau(\tilde{S}')} (w)}{V} \left| \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'} \left(\tilde{S}', \frac{X^{V,\varepsilon}(w)}{V} \right) - \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(\tilde{S}', Z(w)) \right| dw \\
931 &\leq \omega^{\varepsilon,t} \zeta^{\varepsilon,t}.
\end{aligned}$$

933 By summing over the values of the single-molecule trajectories, we also have

$$\begin{aligned}
934 \\
935 \quad \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(Y^{\tilde{S},i}(w), Z(w)) = \tilde{X}_{\tau(\tilde{S})}^V(w) \lambda_{\tilde{S}' + y \rightarrow \tilde{S}'' + y'}(\tilde{S}, Z(w)), \\
936
\end{aligned}$$

937 which implies

$$\begin{aligned}
938 \quad \Upsilon_5 &\leq \sup_{0 \leq u \leq t} \sum_{\tilde{S}'+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} \left\| \frac{\tau(\tilde{S}'')}{\alpha(\tau(\tilde{S}''))} - \frac{\tau(\tilde{S}')}{\alpha(\tau(\tilde{S}'))} \right\|_{\infty} \int_0^u \left| \frac{X_{\tau(\tilde{S})}^{V,\varepsilon}(w)}{V} - \frac{\tilde{X}_{\tau(\tilde{S})}^V(w)}{V} \right| \lambda_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}(\tilde{S}', Z(w)) dw \\
939 &\leq \int_0^t \left\| \frac{X^{V,\varepsilon}(u)}{V} - \frac{\tilde{X}^V(u)}{V} \right\|_{\infty} \hat{\Lambda}_0^u du \\
940 &= \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}^c} \int_0^t \left\| \frac{X^{V,\varepsilon}(u)}{V} - \frac{\tilde{X}^V(u)}{V} \right\|_{\infty} \hat{\Lambda}_0^u du + \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}} \int_0^t \left\| \frac{\hat{X}^{V,\varepsilon}(u)}{V} - \frac{\tilde{X}^V(u)}{V} \right\|_{\infty} \hat{\Lambda}_0^u du. \\
941 &\leq \mathbb{1}_{\mathcal{A}_{V,\varepsilon,t}^c} M^{V,\varepsilon,t} + \int_0^t \left\| \frac{\hat{X}^{V,\varepsilon}(u)}{V} - \frac{\tilde{X}^V(u)}{V} \right\|_{\infty} \hat{\Lambda}_0^u du, \\
942 &
\end{aligned}$$

943 where

$$944 \quad M^{V,\varepsilon,t} = \int_0^t \left(\|Z(u)\|_{\infty} + \varepsilon + \sum_{S \in \mathcal{X}} \alpha(S) \frac{\tilde{X}_S^V(0)}{V} \right) \hat{\Lambda}_0^u du$$

945 is an almost surely finite random variable, non-decreasing in t . Hence, putting every-
946 thing together and applying the Gronwall inequality we have that almost surely

$$\begin{aligned}
947 \quad \sup_{0 \leq t \leq T} \left\| \frac{\pi(\hat{X}^{V,\varepsilon}(t))}{V} - \frac{\tilde{X}^V(t)}{V} \right\|_{\infty} &\leq e^{\hat{\Lambda}_1^T} \frac{\hat{R}}{V} \sup_{0 \leq t \leq T} \left| \bar{U}_1^V \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \right) \right| \\
948 &+ e^{\hat{\Lambda}_1^T} \frac{\hat{r}}{V} \sup_{0 \leq t \leq T} \left| \bar{U}_2^V \left(\sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}'+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \int_0^t \lambda_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}(Y^{\tilde{S},i}(u), Z(u)) du \right) \right| \\
949 &+ e^{\hat{\Lambda}_1^T} \left(\left\| \frac{\pi(X^V(0))}{V} - \frac{\tilde{X}^V(0)}{V} \right\|_{\infty} + \hat{R} \delta_1^{V,\varepsilon,T} + \omega^{\varepsilon,T} \zeta^{\varepsilon,T} + \mathbb{1}_{\mathcal{A}_{V,\varepsilon,T}^c} M^{V,\varepsilon,T} \right). \\
950 &
\end{aligned}$$

951 Now note that if A_1, A_2, \dots, A_j are random variables and a_1, a_2, \dots, a_j are positive
952 real numbers, then

$$953 \quad P \left(\sum_{i=1}^j A_i > \sum_{i=1}^j a_i \right) \leq P \left(\bigcup_{i=1}^j (A_i > a_i) \right) \leq \sum_{i=1}^j P(A_i > a_i).$$

954 Hence, if ν is as in the statement of the theorem and $\nu < \varepsilon$,

$$\begin{aligned}
955 \quad P \left(\sup_{0 \leq t \leq T} \left\| \frac{\pi(X^V(t))}{V} - \frac{\tilde{X}^V(t)}{V} \right\|_{\infty} > \nu \right) &= P \left(\sup_{0 \leq t \leq T} \left\| \frac{\pi(\hat{X}^{V,\varepsilon}(t))}{V} - \frac{\tilde{X}^V(t)}{V} \right\|_{\infty} > \nu \right) \\
956 &\leq P \left(\frac{1}{V} \sup_{0 \leq t \leq T} \left| \bar{U}_1^V \left(\sum_{y \rightarrow y' \in \mathcal{R}} \int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \right) \right| > \nu_1 \right) \\
957 &+ P \left(\frac{1}{V} \sup_{0 \leq t \leq T} \left| \bar{U}_2^V \left(\sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}'+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \int_0^t \lambda_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}(Y^{\tilde{S},i}(u), Z(u)) du \right) \right| > \nu_2 \right) \\
958 &+ p^{V,\varepsilon,T}. \\
959 &
\end{aligned}$$

960 Since for all $t \in [0, T]$

$$961 \quad \int_0^t \lambda_{y \rightarrow y'}^V(X^{V,\varepsilon}(u)) du \leq V \hat{\Lambda}_3^{V,\varepsilon,t}$$

962 and

$$963 \quad \sum_{\tilde{S} \in \tilde{\mathcal{X}} \setminus \{\Delta\}} \sum_{\tilde{S}'+y \rightarrow \tilde{S}''+y' \in \tilde{\mathcal{R}}} \sum_{i=1}^{\tilde{X}_{\tau(\tilde{S})}^V(0)} \int_0^t \lambda_{\tilde{S}'+y \rightarrow \tilde{S}''+y'}(Y^{\tilde{S},i}(u), Z(u)) du \leq V c \hat{\Lambda}_2^t,$$

964 the proof is concluded by Lemma A.1. \square

965 *Proof of Theorem 5.8.* Note that by Lemma 5.7 and by the fact that $\alpha(S) \geq 1$
966 for all $S \in \mathcal{X}$ in (5.6),

$$967 \quad \left\| \frac{\pi(X^V(h))}{V} - \frac{\tilde{X}^V(h)}{V} \right\|_1 \leq \left\| \frac{\pi(X^V(h))}{V} \right\|_1 + \left\| \frac{\tilde{X}^V(h)}{V} \right\|_1$$

$$968 \quad \leq \frac{1}{V} \left(\sum_{S \in \tilde{\mathcal{X}}} \alpha(S) \left(X_S^V(0) + \tilde{X}_S^V(0) \right) \right).$$

969

970 Under the assumption that both $X^V(0)$ and $\tilde{X}^V(0)$ have finite expectation and con-
971 verge in probability to z^* , and by the equivalence of norms in finite dimension, we
972 conclude there exists $M \in \mathbb{R}_{>0}$ such that

$$973 \quad \sup_{V \in \mathbb{Z}_{\geq 1}} E \left[\left\| \frac{\pi(X^V(h))}{V} - \frac{\tilde{X}^V(h)}{V} \right\|_\infty \right] \leq M.$$

974 Hence, if ν is as in Theorem A.5, we have that

$$975 \quad E \left[\sup_{0 \leq t \leq T} \left\| \frac{\pi(X^V(h))}{V} - \frac{\tilde{X}^V(h)}{V} \right\|_\infty \right] \leq \nu + 6M e^{\frac{\hat{\Lambda}_3^{V,\varepsilon,t}}{2} - \frac{\nu_1 \sqrt{V}}{3}}$$

$$976 \quad + 6M e^{\frac{c \hat{\Lambda}_2^t}{2} - \frac{\nu_2 \sqrt{V}}{3}} + MP \left(\left\| \frac{\pi(X^V(0))}{V} - \frac{\tilde{X}^V(0)}{V} \right\|_\infty > \nu_3 \right) + Mp^{V,\varepsilon,T}.$$

977

978

979 The proof is concluded if we can show that for all $T \in \mathbb{R}_{>0}$ and any arbitrary $\eta > 0$,
980 we can fix $\nu_1, \nu_2, \nu_3 \in \mathbb{R}_{>0}$ and $\varepsilon \in (0, m)$ such that $\nu < \eta$ for large enough values
981 of V . Indeed, for any fixed $\varepsilon \in (0, m), T \in \mathbb{R}_{>0}$ the other terms on the right-hand
982 side of the above inequality tend to zero as V goes to infinity. To show that ν can
983 be made smaller than η , simply note that ν_1, ν_2, ν_3 can be chosen as small as desired
984 among the positive real numbers, $\delta_1^{V,\varepsilon,T}$ tends to zero as V goes to infinity for all fixed
985 $\varepsilon \in (0, m)$ by Assumption 3.1, and $\omega^{\varepsilon,T}$ tends to zero as ε tends to zero because the
986 functions $\lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}$ are locally Lipschitz on $\tilde{\mathcal{X}} \times \mathbb{R}_{>0}^d$ by Lemma 5.1. \square

987 **A.4. Proof of Theorem 5.5.** Note that under the assumptions of Theorem 5.5,
988 for all $t \in [0, T]$ $Y^V(t)$ converges in probability to $Y(t)$ by Theorem 5.2. Hence, in
989 order to prove Theorem 5.5, we need to show relative compactness of $\{Y^V\}$ as a
990 sequence of processes with sample paths in $D_{\tilde{\mathcal{X}}}[0, T]$, and conclude by [10, Lemma
991 A2.1]. To prove relative compactness of $\{Y^V\}$, first note that the state space $\tilde{\mathcal{X}}$ is

992 compact. Hence, we only need to show that the jump times do not accumulate as
 993 V tends to infinity, nor tend to 0. Let t_i^V with $i \in \mathbb{Z}_{\geq 1}$ denote the time of the i th
 994 jump of Y^V , let $t_0^V = 0$, and let T^V be the time of the last jump of Y^V in $[0, T]$. Fix
 995 $\delta \in \mathbb{R}_{>0}$ and for all $j \in \mathbb{Z}$ with $-1 \leq j \leq T/\delta$ let $N_j^{V,\delta}$ be the number of jumps of Y^V
 996 in the interval $[j/\delta, \min\{j/\delta + 2\delta, T\}]$. The $N_j^{V,\delta}$ are introduced to control the time
 997 between jumps: whenever two jumps occur at times closer than δ , there necessarily
 998 exists an interval $[j/\delta, \min\{j/\delta + 2\delta, T\}]$ with $j \geq 0$ containing both of them. Also,
 999 whenever the time of a jump is smaller than 0, then $N_{-1}^{V,\delta} \geq 1$. Hence, for all $\varepsilon \in \mathbb{R}_{>0}$
 1000 with $\varepsilon > m$,

$$\begin{aligned}
 1001 \quad P\left(\min_{j=1,\dots,T^V} (t_j^V - t_{j-1}^V) \leq \delta\right) &\leq P\left(N_{-1}^{V,\delta} \geq 1 \text{ or } \max_{j=1,\dots,\lfloor T/\delta \rfloor} N_j^{V,\delta} \geq 2\right) \\
 1002 \quad &\leq P\left(N_{-1}^{V,\delta} \geq 1\right) + \sum_{j=1}^{\lfloor T/\delta \rfloor} P(N_j^{V,\delta} \geq 2) \\
 1003 \quad &\leq P\left(\sup_{0 \leq t \leq T} \left\| \frac{X^V}{V}(t) - Z(t) \right\|_{\infty} > \varepsilon\right) + P(N^\varepsilon(\delta) \geq 1) + \frac{T}{\delta} P(N^\varepsilon(2\delta) \geq 2), \blacksquare \\
 1004
 \end{aligned}$$

1005 where N^ε is a Poisson process with rate

$$1006 \quad B_\varepsilon = \sup_{N \in \mathbb{Z}_{\geq 1}} \sup_{z \in \Omega_1^{\varepsilon,T}} \max_{\tilde{S} \in \tilde{\mathcal{X}}} \sum_{\tilde{S}+y \rightarrow \tilde{S}'+y' \in \tilde{\mathcal{R}}} \lambda_{\tilde{S}+y \rightarrow \tilde{S}'+y'}^V(\tilde{S}, [Vz]),$$

1007 which is finite by Lemma 5.1. Hence, by Theorem 3.2

$$1008 \quad \limsup_{V \rightarrow \infty} P\left(\min_{j=1,\dots,T^V} (t_j^V - t_{j-1}^V) \leq \delta\right) \leq (1 - e^{-\delta B_\varepsilon}) + \frac{T}{\delta} (1 - e^{-2\delta B_\varepsilon} - 2\delta B_\varepsilon e^{-2\delta B_\varepsilon}),$$

1009 which tends to 0 as δ tends to 0. Therefore, $\{Y^V\}$ is relatively compact as a sequence
 1010 of processes with sample paths in $D_{\tilde{\mathcal{X}}}[0, T]$ by [11, Corollary 7.4, Chapter 3], which
 1011 completes the proof.

1012

REFERENCES

- 1013 [1] A. AGAZZI, L. ANDREIS, R. I. PATTERSON, AND D. M. RENGER, *Large deviations for markov*
 1014 *jump processes with uniformly diminishing rates*, Stochastic Processes and their Applica-
 1015 tions, 152 (2022), pp. 533–559.
 1016 [2] A. AGAZZI, A. DEMBO, AND J.-P. ECKMANN, *Large deviations theory for markov jump models of*
 1017 *chemical reaction networks*, The Annals of Applied Probability, 28 (2018), pp. 1821–1855.
 1018 [3] D. F. ANDERSON, D. CAPPELLETTI, J. KIM, AND T. D. NGUYEN, *Tier structure of strongly*
 1019 *endotactic reaction networks*, Stochastic Processes and their Applications, 130 (2020),
 1020 pp. 7218–7259.
 1021 [4] D. F. ANDERSON, A. GANGULY, AND T. G. KURTZ, *Error analysis of tau-leap simulation*
 1022 *methods*, The Annals of Applied Probability, 21 (2011), pp. 2226 – 2262.
 1023 [5] D. F. ANDERSON, D. J. HIGHAM, S. C. LEITE, AND R. J. WILLIAMS, *On constrained langevin*
 1024 *equations and (bio) chemical reaction networks*, Multiscale Modeling & Simulation, 17
 1025 (2019), pp. 1–30.
 1026 [6] D. F. ANDERSON AND T. G. KURTZ, *Stochastic analysis of biochemical systems*, Springer, 2015.
 1027 [7] C. D. BASTIAN AND G. A. REMPALA, *Throwing stones and collecting bones: Looking for*
 1028 *poisson-like random measures*, Mathematical Methods in the Applied Sciences, 43 (2020),
 1029 pp. 4658–4668.
 1030 [8] W. R. K. BUKHSH, C. D. BASTIAN, M. WASCHER, C. KLAUS, S. Y. SAHAI, M. H. WEIR,
 1031 E. KENAH, E. ROOT, J. H. TIEN, AND G. A. REMPALA, *Projecting COVID-19 Cases and*
 1032 *Subsequent Hospital Burden in Ohio*, medRxiv, (2022).

- 1033 [9] F. DI LAURO, W. R. KHUDABUKHSH, I. Z. KISS, E. KENAH, M. JENSEN, AND G. A. REMPALA,
1034 *Dynamic survival analysis for non-markovian epidemic models*, Journal of the Royal So-
1035 ciety Interface, 19 (2022), p. 20220124.
- 1036 [10] P. DONNELLY AND T. G. KURTZ, *A countable representation of the Fleming-Viot measure-*
1037 *valued diffusion*, The Annals of Probability, 24 (1996), pp. 698 – 742.
- 1038 [11] S. N. ETHIER AND T. G. KURTZ, *Markov processes: characterization and convergence*, John
1039 Wiley & Sons Inc, 1986.
- 1040 [12] D. T. GILLESPIE, *Approximate accelerated stochastic simulation of chemically reacting systems*,
1041 The Journal of Chemical Physics, 115 (2001), pp. 1716–1733.
- 1042 [13] H.-W. KANG, T. G. KURTZ, AND L. POPOVIC, *Central limit theorems and diffusion approxima-*
1043 *tions for multiscale markov chain models*, The Annals of Applied Probability, 24 (2014),
1044 pp. 721–759.
- 1045 [14] W. KHUDABUKHSH, S. K. KHALSA, E. KENAH, G. REMPALA, AND J. H. TIEN, *Covid-19 dy-*
1046 *namics in an ohio prison*, medRxiv, (2021).
- 1047 [15] W. R. KHUDABUKHSH, B. CHOI, E. KENAH, AND G. A. REMPALA, *Survival dynamical systems:*
1048 *individual-level survival analysis from population-level epidemic models*, Interface focus,
1049 10 (2020), p. 20190048.
- 1050 [16] T. G. KURTZ, *The relationship between stochastic and deterministic models for chemical reac-*
1051 *tions*, The Journal of Chemical Physics, 57 (1972), pp. 2976–2978.
- 1052 [17] T. G. KURTZ, *Limit theorems and diffusion approximations for density dependent markov*
1053 *chains*, in Stochastic Systems: Modeling, Identification and Optimization, I, Springer,
1054 1976, pp. 67–78.
- 1055 [18] P. MOZGUNOV, M. BECCUTI, A. HORVATH, T. JAKI, R. SIROVICH, AND E. BIBBONA, *A review of*
1056 *the deterministic and diffusion approximations for stochastic chemical reaction networks*,
1057 Reaction Kinetics, Mechanisms and Catalysis, 123 (2018), pp. 289–312.
- 1058 [19] J. R. NORRIS, *Markov chains*, Cambridge university press, 1998.
- 1059 [20] A. PRODHOMME, *Strong gaussian approximation of metastable density-dependent markov*
1060 *chains on large time scales*, arXiv preprint arXiv:2010.06861, (2020).
- 1061 [21] I. SOMEKH, W. R. KHUDABUKHSH, E. D. ROOT, L. K. BOKER, G. REMPALA, E. A. SIMÕES,
1062 AND E. SOMEKH, *Quantifying the population-level effect of the covid-19 mass vaccination*
1063 *campaign in israel: a modeling study*, in Open forum infectious diseases, vol. 9, Oxford
1064 University Press US, 2022, p. ofac087.
- 1065 [22] E. SONTAG, *Structure and stability of certain chemical networks and applications to the kinetic*
1066 *proofreading model of t-cell receptor signal transduction*, IEEE Transactions on Automatic
1067 Control, 46 (2001), pp. 1028 – 1047.
- 1068 [23] M. WASCHER, P. M. SCHNELL, W. R. KHUDABUKHSH, M. QUAM, J. H. TIEN, AND G. A. REM-
1069 PALA, *Monitoring sars-cov-2 transmission and prevalence in populations under repeated*
1070 *testing*, medRxiv, (2021).