

# Mean-Field Limits Beyond Ordinary Differential Equations

Luca Bortolussi, Nicolas Gast

► **To cite this version:**

Luca Bortolussi, Nicolas Gast. Mean-Field Limits Beyond Ordinary Differential Equations. Bernardo, Marco; De Nicola, Rocco; Hillston, Jane. Formal Methods for the Quantitative Evaluation of Collective Adaptive Systems, Programming and Software Engineering, 2016, 978-3-319-34095-1. 10.1007/978-3-319-34096-8\_3. hal-01334358

**HAL Id: hal-01334358**

**<https://hal.inria.fr/hal-01334358>**

Submitted on 20 Jun 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Mean-Field Limits Beyond Ordinary Differential Equations

Luca Bortolussi<sup>1</sup> and Nicolas Gast<sup>2</sup>

<sup>1</sup> DMG, University of Trieste, Trieste, Italy.  
MOSI, Saarland University, Saarbrücken, Germany.  
CNR-ISTI, Pisa, Italy.

<sup>2</sup> Inria, France

**Abstract.** We study the limiting behaviour of stochastic models of populations of interacting agents, as the number of agents goes to infinity. Classical mean-field results have established that this limiting behaviour is described by an ordinary differential equation (ODE) under two conditions: (1) that the dynamics is smooth; and (2) that the population is composed of a finite number of homogeneous sub-populations, each containing a large number of agents. This paper reviews recent work showing what happens if these conditions do not hold. In these cases, it is still possible to exhibit a limiting regime at the price of replacing the ODE by a more complex dynamical system. In the case of non-smooth or uncertain dynamics, the limiting regime is given by a differential inclusion. In the case of multiple population scales, the ODE is replaced by a stochastic hybrid automaton.

**Keywords:** Population models; Markov chain, Mean-field limits; Differential inclusions; Hybrid systems.

## 1 Introduction

Many systems can be effectively described by stochastic population models, for instance biological systems [51], epidemic spreading [1], queuing networks [41]. These systems are composed of a set of objects, agents, or entities interacting together. Each individual agent is typically described in a simple way, as finite state machines with few states. An agent changes state spontaneously or by interacting with other agents in the system. All transitions happen probabilistically and take a random time to be completed. By choosing exponentially distributed times, the resulting stochastic process is a continuous-time Markov chain with a finite state space. Many numerical techniques exist to compute probabilities of such chains [3], and they are part of state-of-the art stochastic model checking tools like PRISM [38] or MRMC [36].

These techniques, however, are limited in their applicability, as they suffer from the state-space explosion: the state-space grows exponentially with the number of agents and even simple agents, when present in large quantities, can generate a huge state space which is far beyond the capabilities of current tools.

This results in the need for approximation techniques to estimate the probabilities and the behaviours of the system. A classic way is to resort to stochastic simulation, which scales better but is still a computationally intensive process for large populations. Precisely in this regime of large populations, mean field analysis offers a viable, often accurate, and much more efficient alternative. The basic idea of mean field is that, when counting the number of agents that are in a given state, the fluctuations due to stochasticity become negligible as the number of agents  $N$  grows. For large  $N$ , the system becomes essentially deterministic.

A series of results, *e.g.*, [4, 7, 37], have established that when the state space of each agent is finite and the dynamics is sufficiently smooth, the system's behaviour converges as  $N$  goes to infinity to a limiting behaviour that is described by system of ordinary differential equations (ODE). The dimension of this system of ODE is equal to the number of states of the individual agents, but independent of the population size  $N$ . The dimension of the differential equation is typically small, hence the numerical integration of these equations is extremely fast. These results show that the intensity of the fluctuations goes to zero as  $1/\sqrt{N}$ . This approach is used in many domains, including computer-based systems [29, 31, 43, 50], epidemic or rumour propagation [17, 34] or bike-sharing systems [24]. Is it also used to construct approximate solutions of stochastic model checking problems [10, 12, 13, 11].

However, these limiting results have two main shortcomings. First, these models cannot deal with discontinuities on the rates of interaction between agents, or uncertainty in model parameters in an obvious way. Second, being able to approximate the number of agents by a continuous variable requires all populations to be large. These limitations are essentially due to restricting the attention to a limiting regime that can be expressed in terms of smooth ODEs.

In this document, we show that by enlarging the set of possible limiting regimes, it is possible to extend the classical framework in multiple directions. We first begin in Section 2 by a concise introduction to classical mean field models and their ODE limits. This section requires basic knowledge of CTMCs and ODEs. We then show in Section 3 how discontinuities and uncertainties can be treated uniformly and consistently considering mean field limits in terms of differential inclusions. We then tackle the presence of multiple population scales in Section 4. We show that when the number of agents in some populations go to infinity while others remain finite, the mean field limit is naturally expressed as a stochastic hybrid automaton, where continuous-deterministic and discrete-stochastic dynamics coexist and modulate each other. Last, we mention other related work and extensions of this framework, for instance to cooperative games, in Section 5.

## 2 The classical mean field framework

In this section, we will introduce the fundamental mean field approximation. We assume the reader familiar with basic concepts of Markov Chains in Continu-

ous Time, see e.g. [44, 23] for an introduction, and with ordinary differential equations.

We start in Section 2.1 by introducing a framework to describe the class of systems amenable of mean field analysis, namely Markov population processes (see also the chapter on spatial representations [26]). We illustrate these concepts in Section 2.2 by means of a classic epidemic spreading model. In Section 2.3, we describe the basic mean-field theorems.

## 2.1 Population Continuous-Time Markov Chains

Population continuous-time Markov chains (PCTMCs) describe a set of interacting agents, which can have different internal states. Interactions involve a small number of agents, and can happen randomly in time, according to an exponential distribution with system-dependent rate. We describe these systems in terms of counting variables and transition classes, following the conventions of [7, 33].

More specifically, a PCTMC  $\mathcal{M}$  model is a tuple  $(\mathbf{X}, \mathcal{T}, \mathbf{x}_0, N)$ , where

- $\mathbf{X} = (X_s)_{s \in \mathcal{S}} \in \mathbb{R}^{|\mathcal{S}|}$  is the population vector. The state space of an agent is  $\mathcal{S}$  and  $X_s \in \mathbb{N}$  counts the number of agents in state  $s \in \mathcal{S}$ . The state space of the model is a finite or countable subset of  $\mathbb{R}^{|\mathcal{S}|}$ .
- $\mathcal{T}$  is the set of transition classes, each of the form  $\eta = (\phi_\eta(\mathbf{X}), \mathbf{v}_\eta, f_\eta(\mathbf{X}))$ , where
  - $\phi_\eta(\mathbf{X}) \in \{0, 1\}$  is a guard predicate, representing a subset of  $S$  in which the transition is active;
  - $\mathbf{v} \in \mathbb{R}^n$  is an update vector, encoding the relative change of  $\mathbf{X}$  induced by the firing of the transition  $\eta$ : the new state will be  $\mathbf{X} + \mathbf{v}_\eta$ ;
  - $f_\eta(\mathbf{X})$  is the rate function, giving the rate at which an  $\eta$  transition is fired as a function of the state space of the system. Typically,  $f_\eta(\mathbf{X})$  is a (locally) Lipschitz continuous function of the population variables.
- $\mathbf{x}_0 \in \mathbb{R}^{|\mathcal{S}|}$  is the initial state of the system.
- $N$  is the population size.

Each PCTMC model  $\mathcal{M}$  defines a CTMC  $\mathbf{X}(t)$  on the state space  $S$ . This chain is characterised by the infinitesimal generator matrix  $Q$  [23], whose off-diagonal entries are given by

$$Q_{\mathbf{x}_1, \mathbf{x}_2} = \sum_{\eta \in \mathcal{T} \text{ s.t. } \mathbf{x}_2 = \mathbf{x}_1 + \mathbf{v}_\eta} \phi_\eta(\mathbf{x}_1) f_\eta(\mathbf{x}_1).$$

An important concept related to population models is the *system size*,  $N$ . Typically, system size is the total (initial) population. In some domains, though, like biochemical networks or ecological models,  $N$  may represent another measure of size, like the volume or the area in which the dynamics described by a PCTMC happens. We refer to [7] for a deeper discussion of this.

## 2.2 Example: SIR epidemic spreading

As a simple and illustrative example, we consider the spreading of a disease in a community of  $N$  agents (which can be humans, animals, computers). The state space of an agent is  $\mathcal{S} = \{S, I, R\}$ . This model is one of the classical examples of a Markov population process and is referred to as the SIR model.

The contagion happens when a susceptible agent ( $X_S$ ) enters in contact with another agent who turns out to be infected (at rate  $k_{si}X_SX_I/N$ ) or enters in contact with an external source of the disease (at rate  $k_iS$ ). Infected individuals spontaneously recover at rate  $k_r$ , and become Recovered ( $X_R$ ) and immune from the disease. This immunity, however, can be lost with rate  $k_s$ .

Formally, the model can be described as a PCTMC with three variables ( $X_S, X_I, X_R$ ), each taking values in the integers  $\{0, \dots, N\}$ , as no birth or death events are considered. The model has four transition classes, all having a guard predicate evaluating to true ( $= 1$ ) in all states :

- Internal infection: ( $\text{true}, \mathbf{e}_I - \mathbf{e}_S, k_{si}X_SX_I/N$ );
- External infection: ( $\text{true}, \mathbf{e}_I - \mathbf{e}_S, k_iX_S$ );
- Recovery: ( $\text{true}, \mathbf{e}_R - \mathbf{e}_I, k_rX_I$ );
- Immunity loss: ( $\text{true}, \mathbf{e}_S - \mathbf{e}_R, k_sX_R$ );

## 2.3 Classic mean field equations

mean field theory answers the following question about population models: what happens when the population is very large? More specifically, it can be shown that, for a large class of models, the dynamics of the system greatly simplifies as the system size goes to infinity. The classic theorem, dating back to the 1970s [37], shows that trajectories of suitably rescaled processes for large populations look deterministic, and in fact converge to the solution of an ordinary differential equation (ODE).

An important operation in the path to mean field is to normalise population processes, dividing variables by the system size, and updating accordingly the transitions. This allows one to compare different models, as they will now have the same scale, intuitively these are population densities. Roughly, the deterministic behaviour appears because fluctuations around the mean of a population process grow as  $\sqrt{N}$ , hence while normalising, i.e. dividing by  $N$  variables, fluctuations will be of magnitude  $1/\sqrt{N}$ , and will thus go to zero.

More formally, consider a population model  $\mathcal{M}^N$ , where we make explicit the dependence on the system size, and define its normalised version  $\hat{\mathcal{M}}^N$  as follows:

- Population variables (and initial conditions) are rescaled by  $N$ :  $\hat{\mathbf{X}}^N = \mathbf{X}/N$ ;
- Transition rates and guard predicates are expressed in the normalised variables, by substituting  $N\hat{\mathbf{X}}^N$  for  $\mathbf{X}$  in the functions:  $\hat{f}^N(\hat{\mathbf{X}}^N) = f(N\hat{\mathbf{X}})$  and  $\hat{\phi}^N(\hat{\mathbf{X}}^N) = \phi(N\hat{\mathbf{X}})$ ;
- Update vectors are rescaled by  $N$ , too:  $\hat{\mathbf{v}}^N = \mathbf{v}/N$ ;

The CTMC associated with the normalised model will be denoted by  $\hat{\mathbf{X}}^N(t)$ .

*Example.* Consider the SIR model of Section 2.2. Its normalised version, for a population of  $N$  agents, has the following four transition classes:

- Internal infection:  $\hat{\mathbf{v}}_{si}^N = (\mathbf{e}_I - \mathbf{e}_S)/N$ ,  $\hat{f}_{si}^N(\hat{\mathbf{X}}) = Nk_{si}\hat{X}_S\hat{X}_I$ ;
- External infection:  $\hat{\mathbf{v}}_i^N = (\mathbf{e}_I - \mathbf{e}_S)/N$ ,  $\hat{f}_i^N(\hat{\mathbf{X}}) = Nk_i\hat{X}_S$ ;
- Recovery:  $\hat{\mathbf{v}}_r^N = (\mathbf{e}_R - \mathbf{e}_I)/N$ ,  $\hat{f}_r^N(\hat{\mathbf{X}}) = Nk_r\hat{X}_I$ ;
- Immunity loss:  $\hat{\mathbf{v}}_s^N = (\mathbf{e}_S - \mathbf{e}_R)/N$ ,  $\hat{f}_s^N(\hat{\mathbf{X}}) = Nk_s\hat{X}_R$ .

As we can see, all transition rates depend linearly on system size. When this happens, rates are called density dependent [7], a condition that usually guarantees the applicability of the mean field results.

*Drift.* The main quantity required to define mean field equations is the drift. The drift is the average direction of change of the population model, conditional on being in a certain state at some time  $t$ . The drift of the normalised model is

$$F(\mathbf{x}) = \sum_{\eta \in \mathcal{T}} \hat{\phi}_\eta(\mathbf{x}) \hat{f}_\eta(\mathbf{x}) \hat{\mathbf{v}}_\eta, \quad (1)$$

Usually, mean field is defined under some additional restrictions on the population model:

- (C1) Guards are true for any  $\mathbf{x} \in S$ , hence the indicator function can be safely removed from the drift:  $F^N(\mathbf{x}) = \sum_{\eta \in \mathcal{T}} \hat{f}_\eta^N(\mathbf{x}) \hat{\mathbf{v}}_\eta^N$ .
- (C2)  $F$  is a Lipschitz continuous function.

Note that by definition of the rescaled model, the drift  $F(\mathbf{x})$  does not depend on  $N$ , because the update vectors are rescaled by  $1/N$  while the transition rates are rescaled by  $N$ . When the drift  $F^N(\mathbf{x})$  does depend on  $N$ , condition (C2) can be replaced in all theorems by a condition (C2'):  $F^N(\mathbf{x})$  converges uniformly as  $N \rightarrow \infty$  to a Lipschitz continuous function  $F(\mathbf{x})$ .

Note that in the SIR model, Conditions (C1) and (C2) are satisfied. For the second one, in particular, we can see that by multiplying a normalised update vector, e.g.  $\hat{\mathbf{v}}_i^N$ , by the corresponding rate, e.g.  $\hat{f}_i^N(\hat{\mathbf{X}}) = Nk_i\hat{X}_S\hat{X}_I$ , the dependency on system size cancels out, so that the drift  $F^N(\mathbf{x})$  is independent of  $N$ . Lipschitz continuity<sup>3</sup> is easily proved. See [7] for a deeper discussion of these conditions. In the following, we will discuss how to weaken these assumptions.

The following theorem can then be proved (see [4, 7, 20]):

**Theorem 1.** *Assume the above conditions C1 and C2 hold and that  $\hat{\mathbf{X}}_0$  converges to  $\mathbf{x}_0$  almost surely (resp. in probability) as  $N$  goes to infinity. Let  $\mathbf{x}$  be the solution of the ODE:*

$$\frac{d}{dt} \mathbf{x}(t) = F(\mathbf{x}(t)) \quad \mathbf{x}(0) = \mathbf{x}_0. \quad (2)$$

Then, for any  $T > 0$ ,

$$\lim_{N \rightarrow \infty} \sup_{t \leq T} \|\hat{\mathbf{X}}^N(t) - \mathbf{x}(t)\| = 0 \quad \text{almost surely (resp. in probability).}$$

<sup>3</sup> In fact, Lipschitz continuity is satisfied only locally, but this is enough for mean field convergence to work.

The theorem essentially states that trajectories of the PCTMC, for large  $N$ , will be indistinguishable from the solution of the mean field ODE restricting to the law of large numbers. An example of the theorem at work, for the SIR model, can be seen in Figure 1. For the SIR model, the mean field approximation is given by the following system of ODEs:

$$\begin{aligned}\dot{x}_S &= -k_i x_S - k_{si} x_S x_I + k_s x_R \\ \dot{x}_I &= k_i x_S + k_{si} x_S x_I - k_r x_I \\ \dot{x}_R &= k_r x_I - k_s x_R\end{aligned}\tag{3}$$

We simulate the model for population size of  $N = 10$ ,  $N = 100$  and  $N = 1000$  agents and report the evolution of the numbers of susceptible or infected agents as a function of time. The parameters are  $k_i = k_{si} = k_s = k_r = 1$  and the initial conditions are  $X_I(0) = X_S(0) = 2X_R(0) = 2N/5$ . Each plot contains three curves: a sample path of one simulation, the mean field (ODE) approximation and an average over  $10^4$  simulations.

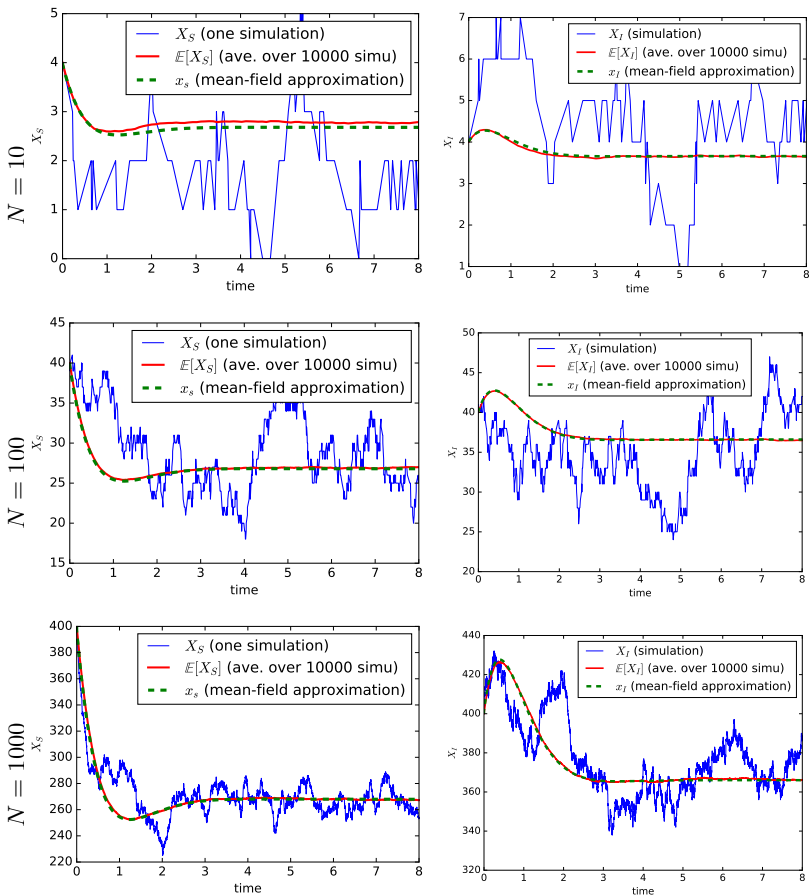
This figure illustrates how large the population size  $N$  has to be for the approximation to be accurate. We observe that the mean field approximation describes correctly the overall dynamics of the PCTMC for  $N = 100$  and  $N = 1000$ . In fact, it can be shown that the rate of convergence in Theorem 1 is of the order of  $1/\sqrt{N}$  [20, 4] but with bounds that often an underestimation of the real convergence speed. In practice, we often observe that the convergence is quicker than this bound. This is particularly true when one considers the average stochastic value:  $\mathbb{E}[\mathbf{X}(t)]$ . To illustrate this fact, we simulated the SIR model  $10^4$  times to compute the values  $\mathbb{E}[\mathbf{X}(t)]$ . We report the evolution of  $\mathbb{E}[X_S(t)]$  and  $\mathbb{E}[X_I(t)]$  with time in Figure 1. We observe than, already for  $N = 10$ , the ODE  $\mathbf{x}(t)$  is extremely close to the value of  $\mathbb{E}[\mathbf{x}(t)]$  computed by simulation. For  $N = 100$  and  $N = 1000$ , the curves are indistinguishable.

Notice that in Theorem 1, the restriction to finite time horizons is fundamental, as convergence at steady state does not necessarily hold. An example is given by the SIR model when  $k_i = 0$ ,  $k_{si} = 3$ ,  $k_s = k_r = 1$  and  $x_R(0) = x_S(0) = x_I(0) = 1/3$ . This initial condition is a fixed point of the mean field ODE which therefore predicts an endemic equilibrium. In the PCTMC model, however, for any  $N$ , the epidemic always extinguishes (i.e. eventually  $X_I(t) = 0$ ) because there is no external infection in this situation. For a deeper discussion of this issue, see the next section as well as [7].

In fact, steady state results can be obtained at the price of adding two additional conditions:

- (C3) For any  $N$ , the PCTMC has a steady-state distribution  $\pi^N$ . The sequence of distributions  $\pi^N$  is tight.<sup>4</sup>
- (C4) The ODE (2) has a unique fixed point  $x^*$  to which all trajectories converge.

<sup>4</sup> A sequence of distributions is  $\pi^N$  tight if their probability does not escape to infinity, i.e. for each  $\epsilon$  there is a compact set  $K$  such that  $\pi^N(K) \geq 1 - \epsilon$  for each  $N$ .



**Fig. 1.** Simulation of the SIR model: comparison between ODE and simulation for various values of  $N$ . We observe that the simulation converges to the ODE as  $N$  goes to infinity. Moreover, even for  $N = 10$ , the average simulation is very close to the ODE.

The condition C3 is a natural condition that is in general true for PCTMC models. For example, if the population CTMC of size  $N$  has a finite number of states, it has a steady-state distribution. Moreover, the tightness of the measure is true if  $\mathbf{X}$  is almost surely bounded.

However, condition C4 is a condition that is often difficult to check for a given set of ODEs. Proving it requires to exhibit a Lyapunov function witnessing global attractiveness of the unique equilibrium point.

**Theorem 2.** *Assume that the above conditions C1, C2, C3 and C4 hold. Then,  $\pi^N$  converges weakly to the Dirac measure  $x^*$ .*

Similarly to Theorem 1, it can be shown that under mild additional conditions, the speed of convergence of the steady-state distribution  $\pi^N$  to  $x^*$  is also



$1/\sqrt{N}$ . For example, it is shown in [53] that this holds when  $x^*$  is exponentially stable, in which case we have

$$\sqrt{\mathbb{E} \left[ \|\mathbf{X}^N(\infty) - x^*\|^2 \right]} = O(1/\sqrt{N}),$$

where  $\mathbf{X}^N(\infty)$  denotes a random point distributed according to the stationary measure  $\pi^N$ .

### 3 Non-continuous dynamics and uncertainties

The classical mean field models make the assumption that the drift  $F$ , given by (1), is Lipschitz-continuous. Yet, this is not the case in many practical problems. For example, this occurs when a transition  $\eta$  has a guard predicate  $\phi_\eta$  that is true only in a sub-part of the domain. This causes a discontinuity in the drift between the two regions where the predicate is true or false. In this case, the ODE  $\dot{x} = F(x)$  is often not well-defined. A classical way of overcoming this difficulty is to replace the ODE by a differential inclusion  $\dot{x} \in F(x)$ .

In this section, we give some general discussion of differential inclusions. We then show how the classical mean field results can be generalised to differential inclusion dynamics in the case of discontinuous dynamics. Last, we show in Section 3.3 how this framework can be used to deal with uncertainties in the parameters. This section collects results from [6, 28].

#### 3.1 The differential inclusion limit

Let  $\mathcal{M} = (\mathbf{X}, \mathcal{T}, \mathbf{x}_0, N)$  be a population model as defined in Section 2.1. Each transition class has the form  $\eta = (\phi(\mathbf{X}), \mathbf{v}_\eta, f_\eta(\mathbf{X}))$ . Recall that the drift of the stochastic system, defined in Equation (1), is

$$F(\mathbf{x}) = \sum_{\eta \in \mathcal{T}} \hat{\phi}_\eta(\mathbf{x}) \hat{f}_\eta(\mathbf{x}) \hat{\mathbf{v}}_\eta, \quad (4)$$

The classical mean field results presented in the previous section (Theorems 1 and 2) apply when the guard predicates  $\phi_\eta$  always evaluate to true. This condition guarantees that, if the functions  $f_\eta$  are Lipschitz continuous, the function  $F$  is also Lipschitz continuous. This ensures that the ODE  $\dot{x} = F(x)$  is well-defined: from any initial condition, it has a unique solution. However, this no longer holds when guard predicates can take the two values true and false. In this case, the drift  $F$  is not continuous and Theorems 1 and 2 no longer apply.

One of the reasons for the inapplicability of those theorems is that when  $F$  is not a Lipschitz-continuous function, the ODE  $\dot{x} = F(x)$  does not necessarily have a solution. For example, let  $F : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $F(x) = -1$  if  $x \geq 0$  and  $F(x) = 1$  if  $x < 0$ . The ODE  $\dot{x} = F(x)$  starting in 0 has no solution. A natural way to overcome this limitation is to use differential inclusions.

Let  $G$  be a multivalued map, that associates to each  $x \in \mathbb{R}^S$  a set  $G(x) \subset \mathbb{R}^S$ . A trajectory  $\mathbf{x}$  is said to be a solution of the differential inclusion  $\dot{x} \in G(x)$  starting in  $x_0$  if:

$$x(t) = x_0 + \int_0^t g(s) ds, \quad \text{where for all } s: g(s) \in G(x(s)).$$

The sufficient condition for the existence of at least one solution  $\mathbf{x} : [0, \infty) \rightarrow \mathbb{R}^d$  of a differential inclusion from any initial condition  $x_0$  is the following (see [2]):

(C5) For all  $x \in \mathbb{R}^d$ ,  $G(x)$  is closed, convex and non-empty;  $\sup_{x \in \mathbb{R}^d} |G(x)| < \infty$  and  $G$  is upper-hemicontinuous.<sup>5</sup>

**Theorem 3 ([28]).** *Let  $\mathcal{M} = (\mathbf{X}^N, \mathcal{T}, \mathbf{x}_0, N)$  be a population model with drift  $F$  such that there exists a function  $G$  satisfying (C5) such that for all  $x \in \mathbb{R}^d$ :  $F(x) \in G(x)$ . Let  $\mathcal{S}$  be the set of solutions of the differential inclusion  $\dot{x} \in G(x)$  starting in  $x_0$ . Then, for all  $T$ , almost surely*

$$\lim_{N \rightarrow \infty} \inf_{\mathbf{x} \in \mathcal{S}_{x_0}} \sup_{t \in [0, T]} \|\mathbf{X}^N(t) - x(t)\| = 0.$$

In other words, as  $N$  grows, the distance between the stochastic process  $\mathbf{X}$  and the set of solutions of the differential inclusion goes to 0. If this set has a unique solution, then  $\mathbf{X}^N$  converges to this solution.

Theorem 3 is a generalisation of Theorem 1 that relaxes the condition (C1) by using a larger drift function  $G$ . We will see in the next section a natural way to choose  $G$  when the drift is not continuous. The price to be paid by this generalisation is composed of two drawbacks. First, differential inclusions can have multiple solutions. Theorem 3 implies that  $\mathbf{X}^N$  gets closer to the solutions of the differential inclusion but does not indicate towards which solutions the process will converge. Second, the speed of convergence of  $\mathbf{X}^N$  to  $\mathcal{S}$  is unknown, apart in the special case of one-sided Lipschitz drift, for which the distance between  $\mathbf{X}^N$  and  $\mathcal{S}$  decays in  $1/\sqrt{N}$  (see [28]). In general, the convergence appears to be slower in the case of non-continuous dynamics (see Figure 2).

The steady-state of a non-continuous PCTMC can also be approximated by using the same approximation. In fact, the results of Theorem 2 can also be directly generalised to the case of non-continuous dynamics: if the differential inclusion has a unique point  $x^*$  to which all trajectories converge, then the steady-state distribution of  $\mathbf{X}^N$  concentrates on  $x^*$  as  $N$  goes to infinity (see [6, 28]).

<sup>5</sup>  $G$  is upper-hemicontinuous if for all  $x, y \in \mathbb{R}^d$ ,  $x_n \in \mathbb{R}^d$ ,  $y_n \in F(x_n)$ ,  $\lim_{n \rightarrow \infty} x_n = x$  and  $\lim_{n \rightarrow \infty} y_n = y$ , then  $y \in F(x)$ .

### 3.2 Application to discontinuous dynamics

A natural way to define a multivalued map that satisfies (C5) is to consider the multivalued map  $\bar{F}$ , defined by

$$\bar{F}(x) = \bigcap_{\varepsilon > 0} \text{convex\_hull} \left( \bigcup_{x': \|x-x'\| \leq \varepsilon} F(x') \right). \quad (5)$$

It is shown in [28] that if  $F$  is bounded, then  $\bar{F}$  satisfies (C5). By Theorem 3, this implies that, regardless of the properties of the original drift  $F$ , the trajectories of the stochastic system  $\mathbf{X}^N$  converge to the solution of the differential inclusion  $\dot{x} \in \bar{F}(x)$ .

When  $F$  is continuous at a point  $x \in \mathbb{R}^d$ , the  $\bar{F}(x) = F(x)$ . When  $F$  is not continuous in  $x \in \mathbb{R}^d$ ,  $\bar{F}(x)$  is multivalued. To give a concrete example, let us consider the SIR model of Section (2.2) in which we add an additional transition corresponding to the treatment of some infected people. This treatment is applied when the proportion of infected people is greater than 0.3 and changes an infected person into a susceptible individual at rate  $k_t$ . This corresponds to a transition class:

- Treatment:  $\eta = (\mathbf{1}_{X_I \geq 0.3}, \mathbf{e}_S - \mathbf{e}_I, k_t X_I)$

Adding this transition to the original ODE (3), the drift is given by

$$F(x) = \begin{pmatrix} -k_i x_S - k_{si} x_S x_I + k_s x_R + k_t \mathbf{1}_{x_I \geq 0.3} x_I \\ k_i x_S + k_{si} x_S x_I - k_r x_I - k_t \mathbf{1}_{x_I \geq 0.3} x_I \\ k_r x_I - k_s x_R \end{pmatrix}, \quad (6)$$

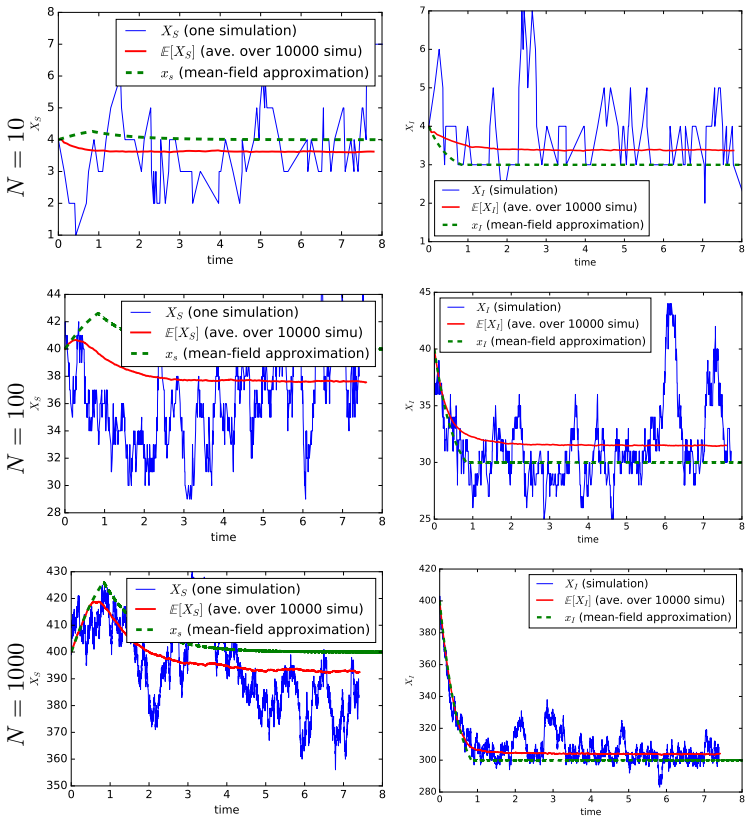
where the guard predicate leads to the term  $\mathbf{1}_{x_I \geq 0.3}$ .

This drift is not continuous in  $x$ . In fact, it can be shown that because of this discontinuity, the corresponding ODE has no solution on  $[0, \infty)$  starting from  $x_0 = (.4, .4, .2)$ . The corresponding  $\bar{F}$  defined by Equation (5) is then given by

$$\bar{F}(x) = \begin{pmatrix} -k_i x_S - k_{si} x_S x_I + k_s x_R + k_t \mathbf{1}_{x_I > 0.3} x_I + k_t [0, x_I] \mathbf{1}_{x_I = 0.3} \\ k_i x_S + k_{si} x_S x_I - k_r x_I - k_t \mathbf{1}_{x_I > 0.3} x_I - k_t \mathbf{1}_{x_I = 0.3} [0, x_I] \\ k_r x_I - k_s x_R \end{pmatrix},$$

where the notation  $a + [b, c]$  denotes the set  $[a + b, a + c]$ .

It can be shown that the differential inclusion  $\dot{x} \in \bar{F}(x)$  has a unique solution,  $\mathbf{x}$ . Hence, Theorem 3 applies to show that  $\mathbf{X}^N$  converges to  $\mathbf{x}$  as  $N$  goes to infinity. To illustrate this fact, we simulated the modified SIR model with the treatment policy and report the results in Figure 2. We observe that, as stated by Theorem 3,  $\mathbf{X}^N$  converges to  $\mathbf{x}$  as  $N$  goes to infinity. In this case, the convergence appears to be slower than for the Lipschitz-continuous case. This is especially visible when looking at the average values  $\mathbb{E}[X_S]$  and  $\mathbb{E}[X_I]$ : in Figure 1, we observe that for the Lipschitz-continuous case,  $\mathbb{E}[X_S]$  is almost equal to  $x_s$  already for  $N = 10$ . In the non-continuous case, Figure 2 indicates that  $\mathbb{E}[X_I]$  does converge to  $x_I$  but at a much slower rate.



**Fig. 2.** Simulation of the SIR model: comparison between differential inclusions and simulation for various values of  $N$ . We observe that the simulation converges to the solution of the differential inclusion as  $N$  goes to infinity. Moreover, even for  $N = 10$ , the average simulation is very close to the solution of the differential inclusion.

### 3.3 Imprecise and uncertain models

Stochastic models are one way of representing uncertainties in a system but they depend on parameters whose precise values are not always known. The differential inclusion framework is also well adapted to study models with imprecise or unknown parameters. Following [6], we distinguish two ways to model uncertainties in models of complex systems:

- **Imprecise scenario:** Some parameters  $\vartheta$  can depend on features of the environment external to the model. We fix a set  $\Theta$  of possible values for  $\vartheta$  and assume that  $\vartheta$  depends on time  $t$  and can take any value of  $\Theta$  at any time instant, i.e. that  $\vartheta_t \in \Theta$ .
- **Uncertain scenario:** In a simpler scenario, a parameter  $\vartheta$  is assumed fixed, but its precise value not known precisely. In this case, we just assume that  $\vartheta \in \Theta$ , where  $\Theta$  is the possible set of values of  $\vartheta$ , as above.

An imprecise or uncertain PCTMC model is a tuple  $(\mathbf{X}, \mathcal{T}, x_0, \Theta, N)$ , where  $\Theta$  is a set of parameters. The difference with classical PCTMC is that the rate function  $f_\eta(\mathbf{X}, \theta)$  of each transition class  $\eta \in \mathcal{T}$  depends on a parameter  $\theta \in \Theta$ .

The differential inclusion framework can be used to study the limits of imprecise and uncertain PCTMC. For the uncertain scenario, there is a differential inclusion  $\dot{x} \in F(x, \vartheta)$  associated with each parameter  $\vartheta$ . Denoting by  $S_\vartheta$  the set of solutions of this differential inclusion, Theorem 3 shows that, as  $N$  goes to infinity, any sequence of uncertain trajectories  $\mathbf{X}^N$  converges to the set  $S_{\text{uncertain}} = \bigcup_{\vartheta \in \Theta} S_\vartheta$ . The differential inclusion corresponding to the imprecise scenario is  $\dot{x} \in \bigcup_{\vartheta \in \Theta} F(x, \vartheta)$ . Denoting by  $S_{\text{imprecise}}$  the set of solutions of this differential inclusion, Theorem 3 shows that, as  $N$  goes to infinity, any imprecise trajectory  $\mathbf{X}^N$  converges to  $S_{\text{imprecise}}$ .

Some numerical methods are developed in [6] to compute or approximate the set of solutions of differential inclusions corresponding to the imprecise and the uncertain model. In particular, we obtain the most precise results by describing the set of reachable values of  $x$  at time  $t$  as a maximisation problem. Then Pontryagin's maximum principle [46] can be used to numerically compute the solution.

## 4 Hybrid mean field limits

In the previous sections, we considered scenarios where all populations of the model are large and grow with the system size. This allows one to prove that their density has vanishing fluctuations around the mean, given by the solution of the mean field equation. However, in many practical cases, there may be multiple population scales in a model, typically in the form of some entities being present in small numbers, independent of the total population size. Examples can be found in genetic regulatory networks, where genes are present in a fixed quantity, typically one or few copies, or more generally in the presence of a centralised form of control [8]. This suggests that in these scenarios we need to consider mean field models in which the continuous and deterministic limit dynamics of parts of the system coexists with the discrete and stochastic dynamics of other parts. Mathematically, this behaviour is captured by stochastic hybrid systems (SHS) [8, 21], which will be introduced in the next subsection.

### 4.1 Stochastic Hybrid Systems

We introduce a model of SHS, essentially borrowing from the treatment of [8, 15] of a class of stochastic hybrid processes known as Piecewise-Deterministic Markov Processes (PDMP) [21].

We consider two sets of variables, the *discrete variables*  $\mathbf{Z} = Z_1, \dots, Z_k$  and the *continuous variables*  $\mathbf{Y} = Y_1, \dots, Y_m$ . The former describes populations that remain discrete also in the mean-field limit, while the second describes populations that will be approximated as continuous. We call  $E = E_d \times E_c$  the hybrid state space of the SHS, with  $E_d \subset \mathbb{N}^k$  a countable set of possible values

for  $\mathbf{Z}$ , and with  $E_c \subset \mathbb{R}^m$  the continuous state space in which variables  $\mathbf{Y}$  can take values. Each possible value that the vector  $\mathbf{Z}$  can take is called a discrete mode, and can be identified with a node in a graph describing the transitions of the discrete states. This graph-based point of view is taken in the definition of stochastic hybrid automata, see e.g. [15].

The evolution of the continuous state is governed by an  $m$ -dimensional *vector field*  $F(\mathbf{Z}, \mathbf{Y})$ , depending on the continuous and the discrete variables. Hence, the continuous variables will evolve following the solution of the differential equation defined by  $F$ , which can be different in each discrete mode  $\mathbf{z}$ . Such a mode-specific continuous dynamics is one of the characteristic features of SHS.

The dynamics of the discrete state is governed by a stochastic Markovian dynamics, specified by two quantities: a *rate function*  $\lambda(\mathbf{Z}, \mathbf{Y})$ , depending both on discrete and continuous variables, and a *jump or reset kernel*  $R(\mathbf{Z}, \mathbf{Y}, \cdot)$ , specifying for each  $\mathbf{Z}, \mathbf{Y}$  a distribution on  $E$ , giving the state in which the system will find itself after a jump of the discrete transition. For the purpose of this chapter, we can restrict ourselves to finitely supported reset kernels, defined by a finite set of pairs of update vectors  $\{(\mathbf{v}_j^d, \mathbf{v}_j^c) \mid j = 1, \dots, h\}$  and associated probability functions  $p_j(\mathbf{Z}, \mathbf{Y})$ , giving the likelihood of jumping from state  $\mathbf{Z}, \mathbf{Y}$  to state  $\mathbf{Z} + \mathbf{v}_j^d, \mathbf{Y} + \mathbf{v}_j^c$ , if a stochastic event fires when the system is in state  $\mathbf{Z}, \mathbf{Y}$ . This, in turn, happens after an exponentially distributed delay with rate  $\lambda(\mathbf{Z}, \mathbf{Y})$ .

Discrete and continuous dynamics in a SHS are intertwined. The system starts in a given state  $\mathbf{z}_0, \mathbf{y}_0$ , and its continuous state evolves following the solution of the initial value problem  $\frac{d}{dt} \mathbf{Y}(t) = F(\mathbf{z}_0, \mathbf{Y}(t))$ ,  $\mathbf{Y}(0) = \mathbf{y}_0$ . This continuous flow will go on until a discrete event will happen, at a random time governed by an exponential distribution with rate  $\lambda(\mathbf{Z}, \mathbf{Y}(t))$ . Note that, as  $\mathbf{Y}$  will change value in time following the flow of the vector field, the rate of a discrete jump is also time-dependent. When a discrete transition happens, say in state  $\mathbf{z}, \mathbf{y}$ , then the system will jump to the state  $\mathbf{z} + \mathbf{v}_j^d, \mathbf{y} + \mathbf{v}_j^c$  with probability  $p_j(\mathbf{Z}, \mathbf{Y})$ . Note that both the discrete mode and the value of continuous variables can change. From this new state, the system continues to evolve following the dynamics given by the vector field in the new discrete mode. The overall dynamics is given by an alternation of periods of continuous evolution interleaved by discrete jumps. For a proper mathematical formalisation of this process, we refer the interested reader to [8, 21].

## 4.2 From PCTMC to SHS

In this section we will show how to construct a SHS from a PCTMC  $(\mathbf{X}, \mathcal{T}, \mathbf{x}_0, N)$ , and how to guarantee the asymptotic correctness of the method. The starting point is a partition of the variables  $\mathbf{X}$  of the PCTMC into two distinct classes: discrete and continuous. We will denote discrete variables with  $\mathbf{Z}$  and continuous ones with  $\mathbf{Y}$ , so that  $\mathbf{X} = \mathbf{Z}, \mathbf{Y}$ . Transitions  $\mathcal{T}$  also have to be separated in two classes: discrete  $\mathcal{T}_d$  and continuous  $\mathcal{T}_c$ . Intuitively, continuous transitions and variables will define the continuous dynamics, and discrete transitions and variables the discrete one. The only request is that continuous transitions do not

affect discrete variables, i.e. for each  $\eta \in \mathcal{T}_c$ ,  $\mathbf{v}_\eta[\mathbf{Z}] = \mathbf{0}$ , where  $\mathbf{v}_\eta[\mathbf{Z}]$  denotes the vector  $\mathbf{v}_\eta$  restricted to the components of  $\mathbf{Z}$ .

*Remark 1.* The choice of how to partition variables and transitions into discrete and continuous is not obvious, and depends on the model under consideration. Often, this is easily deduced from model structure, e.g. due to the presence of conservation laws with a small number of conserved agents. A further help comes from the request to make explicit in the rates and updates the dependency on system size  $N$ . An alternative is to define rules to automatically switch between a discrete or a continuous representation of variables and transitions, depending on the current state of the model. We refer the interested reader to the discussion in [8] for further details.

*Normalisation of continuous variables.* To properly formalise hybrid mean-field limits, we need to perform a normalisation operation on the continuous variables, taking system size into account (hence we will use the superscript  $N$  from now on). This can be obtained as in Section 2.3, by introducing the normalised variables  $\hat{\mathbf{Y}}^N = \mathbf{Y}/N$ , and expressing rates, guards, and update vectors with respect to these normalised variables. Note that normalised update vectors  $\hat{\mathbf{v}}_\eta^N$  are divided by  $N$  only in the continuous components, as the discrete variables are not rescaled. Transitions, after normalisation, have to satisfy some scaling conditions:

- (Tc) Continuous transitions  $\eta \in \mathcal{T}_c$  are such that  $\hat{f}_\eta^N(\mathbf{Z}, \hat{\mathbf{Y}})/N \rightarrow f_\eta(\mathbf{Z}, \hat{\mathbf{Y}})$ , as  $N \rightarrow \infty$ , uniformly on  $\hat{\mathbf{Y}}$  for each  $\mathbf{Z}$ . The limit function is required to be (locally) Lipschitz continuous. Furthermore, their non-normalised update is independent of  $N$ . Guards can depend only on discrete variables:  $\phi_\eta = \phi_\eta(\mathbf{Z})$ .
- (Td) Discrete transitions  $\eta \in \mathcal{T}_d$  are such that their rate function  $\hat{f}_\eta^N(\mathbf{Z}, \hat{\mathbf{Y}})$  converges (uniformly in  $\hat{\mathbf{Y}}$  for each  $\mathbf{Z}$ ) to a continuous function  $f_\eta(\mathbf{Z}, \hat{\mathbf{Y}})$ . Their normalised jump vector  $\hat{\mathbf{v}}_\eta^N$  has also to converge to a vector  $\hat{\mathbf{v}}_\eta$  as  $N$  diverges. Guards can depend only on discrete variables:  $\phi_\eta = \phi_\eta(\mathbf{Z})$ .

Note that, for discrete transitions, we consider the change in the normalised continuous variables, and we admit that the update vectors can depend on  $N$ . In particular, the update for continuous variables can be linear in  $N$ , thus resulting in a non-vanishing jump in the density, in the large  $N$  limit. This means that limit discrete transitions may also induce jumps on continuous variables.

*Construction of the limit SHS.* Given a family of PCTMC models  $(\mathbf{X}, \mathcal{T}, \mathbf{x}_0, N)$ , indexed by  $N$ , with a partition of variables into  $\mathbf{Z}, \mathbf{Y}$  and transitions into  $\mathcal{T}_c, \mathcal{T}_d$  we can normalise continuous variables and formally define the SHS associated with it:

- The vector field defining the continuous dynamics of the SHS is given by the following drift:

$$F(\mathbf{Z}, \hat{\mathbf{Y}}) = \sum_{\eta \in \mathcal{T}_c} \mathbf{v}_\eta I\{\phi_\eta(\mathbf{Z})\} f_\eta(\mathbf{Z}, \hat{\mathbf{Y}})$$

- The jump rate of the SHS is given by

$$\lambda(\mathbf{Z}, \hat{\mathbf{Y}}) = \sum_{\eta \in \mathcal{T}_d} I\{\phi_\eta(\mathbf{Z})\} f_\eta(\mathbf{Z}, \hat{\mathbf{Y}})$$

- The reset kernel is specified by the pair of update vectors  $(\hat{\mathbf{v}}_\eta[\mathbf{Z}], \hat{\mathbf{v}}_\eta[\hat{\mathbf{Y}}])$  and by the probability

$$p_\eta(\mathbf{Z}, \hat{\mathbf{Y}}) = \frac{I\{\phi_\eta(\mathbf{Z})\} f_\eta(\mathbf{Z}, \hat{\mathbf{Y}})}{\lambda(\mathbf{Z}, \hat{\mathbf{Y}})},$$

for each transition  $\eta \in \mathcal{T}_d$ .

- The initial state is  $\mathbf{z}_0, \hat{\mathbf{y}}_0$ .

*SIR model with vaccination.* We consider now an extension of the SIR model of Section 2.2, with the possibility of starting a vaccination campaign of susceptible individuals. The model has an additional variable,  $X_V \in \{0, 1\}$ , which is going to be the only discrete variable of the system and encodes if the vaccination is in force or not. We further have two additional transitions:

- Vaccination of susceptible:  $(\mathbf{true}, \mathbf{e}_R - \mathbf{e}_S, k_v X_S X_V)$ ;
- Activation of the vaccination policy:  $(X_V = 0, \mathbf{e}_V, k_a X_I / N)$ ;

The first transition, which will be a continuous transition, models the effect of vaccination, moving agents from susceptible to recovered state. Note that the rate depends on  $X_V$ , hence the transition is in force only if  $X_V = 1$ . We could alternatively specify the same behaviour by introducing a guard in the transition, depending only on discrete variables, which would result in a rate active only in a subset of discrete modes. The second transition will be kept discrete and model the activation of the vaccination policy. Its rate depends on the density of infected individuals (the higher the infected, the higher the activation rate). The guard on  $X_V$  allows the activation transition to be in force only when the vaccination is inactive.

In Figure 3 (left), we show a trajectory of the system for  $N = 100$ , and compare it with a trajectory of the limit SHS. Parameters are  $k_{si} = 0$ ,  $k_i = 1$ ,  $k_r = 0.1$ ,  $k_s = 0.01$ ,  $k_v = 2$ ,  $k_a = 2$ ,  $X_S(0) = 0.95N$ ,  $X_I(0) = 0.05N$ ,  $X_R(0) = 0$ . As we can see, around time  $t = 2.5$  there is a sudden drop on the number of susceptible individuals, caused by the beginning of vaccination. The similarity between the PCTMC and the SHS trajectories is a clear hint on the existence of an underlying convergence result.

*Mean-field convergence results.* Consider a family of PCTMC model  $(\mathbf{X}, \mathcal{T}, \mathbf{x}_0, N)$ , and denote by  $(\mathbf{Z}^N(t), \hat{\mathbf{Y}}^N(t))$  the normalised CTMC associated with it, for size  $N$ , where we made explicit the partition of variables into discrete and continuous. Denote by  $(\mathbf{z}(t), \mathbf{y}(t))$  the limit SHS, constructed according to the recipe of this section. We can then prove [8] the following theorem:



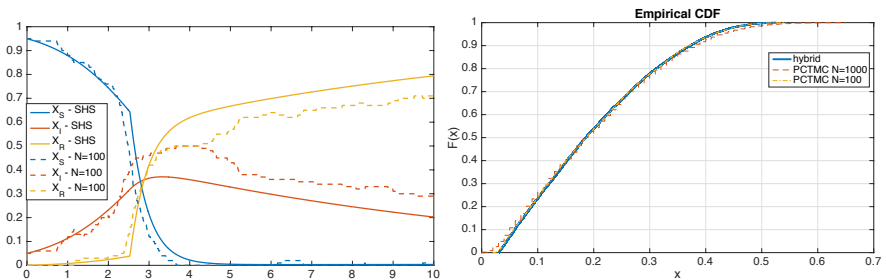
**Theorem 4.** Assume  $(\mathbf{z}_0^N, \hat{\mathbf{y}}_0^N) \rightarrow (\mathbf{z}_0, \mathbf{y}_0)$  and that the transitions of the PCTMC model satisfy conditions (Tc) and (Td). Then

$$(\mathbf{Z}^N(t), \hat{\mathbf{Y}}^N(t)) \Rightarrow (\mathbf{z}(t), \mathbf{y}(t)),$$

for all times  $t \geq 0$ , where  $\Rightarrow$  denotes weak convergence.<sup>6</sup>

This theorem states that the distribution of the PCTMC will look like the distribution of the SHS for large  $N$ . In particular, if  $E_d$  is finite (i.e. there is a finite number of discrete modes) and  $E_c$  is compact, then all conditional and unconditional moments of the distribution converge.

In Figure 3 (right), we see the theorem at work in the epidemic with vaccination example. The figure compares the empirical cumulative distribution of the density of infected individuals at time  $t = 10$ . The curves look quite similar already for  $N = 100$ , and are almost identical for  $N = 1000$ .



**Fig. 3.** Left: comparison of a simulated trajectory of the PCTMC model, for  $N = 100$ , with a simulated trajectory of the limit SHS. Right: comparison of the empirical cumulative distribution of  $\hat{X}_I$  at time  $t = 10$  of the limit SHS and the PCTMC models for  $N = 100$  and  $N = 1000$ .

### 4.3 Extensions of the hybrid limit framework

The hybrid mean field limit presented in the previous section can be extended in many ways, as discussed in [8]. Here we will sketch them briefly, referring the interested reader to [8] for further details.

One possible direction to enrich the framework is to consider forced transitions. In the context of SHS, these are discrete jumps happening as soon as a condition on the system variables becomes true. Typically, they are introduced by constraining the continuous state space  $E_c$  (in a mode dependent way), and forcing a jump to happen as soon as the trajectory of the continuous variables hits the boundary  $\partial E_c$  [21]. Then a jump is done according to the reset kernel

<sup>6</sup> In fact, weak convergence holds for  $(\mathbf{z}, \mathbf{y})$  as processes in the Skorokhod space of cadlag functions, see [8]. For a definition of weak convergence, see [5].

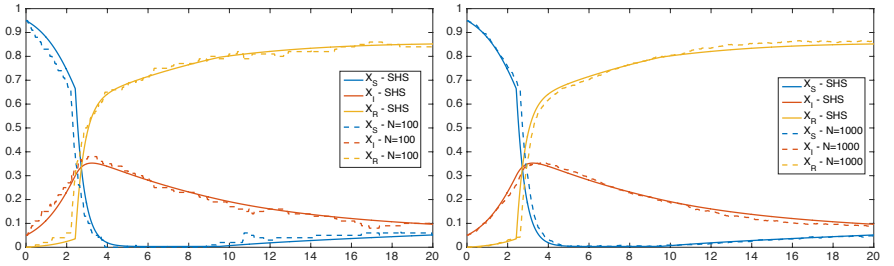
$R$ , whose definition has to be extended on the boundary  $\partial E_c$ . Hence, discrete jumps may happen at stochastic times, or when the condition for a forced jump is met.

In the PCTMC setting, introducing forced transitions requires us to allow transitions with an infinite rate and with a non-trivial guard, firing as soon as their guard becomes true. Their guards, then, can be used to constrain the continuous state space. Hence,  $E_c$  will be defined in each mode as the interior of the complement of the region obtained by taking the union of the satisfaction sets of all the guards of forced transitions. The reset kernel in a point of the boundary  $\partial E_c$  will then be defined by the active immediate transitions at that point, choosing uniformly among the active transitions.<sup>7</sup>

As an example, consider again the SIR model with vaccination, but assume its activation and deactivation is threshold-based: when the density of infected becomes greater than a threshold  $I_{high}$ , the vaccination is started, while if it falls below  $I_{low}$ , the vaccination is stopped. In the PCTMC model, we would have the following two transitions in place of the stochastic one discussed previously:

- Activation of the vaccination policy:  $(X_I \geq NI_{high}, \mathbf{e}_V, \infty)$ ;
- Deactivation of the vaccination policy:  $(X_I \leq NI_{low}, -\mathbf{e}_V, \infty)$ .

Theorem 4 is readily extended to the presence of instantaneous transitions, under some additional technical conditions on the vector fields (called transversal crossing), see [8] for details. The validity of the result for the SIR model with vaccination is illustrated in Figure 4, where we compare simulations of the PCTMC model and the limit SHS, for thresholds  $I_{high} = 0.3$  and  $I_{low} = 0.2$ .



**Fig. 4.** Left: comparison of a simulated trajectory of the PCTMC model with instantaneous transitions (normalised variables), for  $N = 100$  (left) and  $N = 1000$  (right), with a simulated trajectory of the limit SHS.

Other extensions of the hybrid mean field include dealing with guards in deterministic and stochastic transitions. These introduce discontinuities in the

<sup>7</sup> In [8], weights are introduced to solve non-determinism between instantaneous transitions. Furthermore, the possibility of seeing a chain of instantaneous events firing is taken into account. Termination of this chain is discussed in [15] (where it is proved undecidable for countable state spaces), and in [27], where sufficient and testable conditions for termination are given.

model, and require further technical conditions for the limit theorems to hold. As for guards in the continuous transitions, this in fact requires one to introduce in the hybrid context the mean field techniques based on differential inclusions of Section 3.

## 5 Related work and examples

*Load balancing and discontinuous dynamics.* The use of mean field approximation is popular for studying load-balancing policies in server farms. In such a system, an object represents a server and its state is typically the number of jobs that are waiting in its queue. A popular randomised load balancing policy is the *two-choice* policy: for each incoming packet two servers are picked at random and the job is allocated to the least-loaded of the two. This policy has been successfully analysed by classical mean field techniques in [43] where it is shown that it leads to an important gain of performance compared to a purely random allocation. The classical approach then fails when one considers a centralised load balancing policy such as *join the shortest queue* because these policies lead to discontinuous dynamics. As demonstrated in [28, 49, 24], these problems can be modelled and resolved by using differential inclusions.

*Heterogeneous systems and uncertainties.* Another problem where differential inclusions can help is the case of heterogeneous systems. In such cases, there is a large number of objects each having distinct parameters. One example is the caching model of [31] in which an object  $i$  has a popularity  $p_i$ . One possibility to solve the problem is to consider a system of  $N \times S$  ODEs, where  $N$  is the number of objects and  $S$  the dimension of the state space. This method scales linearly in the number of objects but might still be problematic for large populations. An alternative is to consider upper and lower bounds on the dynamics [48] or to study a PDE approximation when the number of objects is large, see for example [25].

*Hybrid mean field limits.* The use of hybrid approximation of population models is quite common in the area of systems and synthetic biology, where genes and often mRNA molecules are present in such low numbers (genes usually in one copy) that classic mean field assumptions are not correct and can lead to models failing to capture important features of the system like bursting protein expression [39]. In addition to the investigation of hybrid limits, carried out independently for general population models [9, 8] and more specifically for gene networks [19], considerable work has been done in hybrid simulation [45] and in developing moment closure techniques for hybrid limits [32].

*Mean field games.* Game theory studies the decisions taken by competing rational agents. Recently, the notion of mean field games has been introduced in [40, 35] to model decisions in systems composed of a large number of agents. In a mean field game, each agent tries to minimise an objective function that

depends on the average behaviour of the population of agents but not on the action of a precise agent. This simplifies the analysis of Nash equilibria that are replaced by mean field equilibria. This theory is used for modelling purposes but also to solve optimisation problems in a decentralised way [52, 30]. It can be shown that in certain cases, mean field equilibria are the limit of a sub-class of Nash equilibria [22].

*Modelling languages supporting a mean field semantics.* In the last ten years, there has been a considerable interest in extending stochastic modelling languages, in particular stochastic process algebras, in order to derive automatically mean field equations. Examples are e.g. [47, 18, 14]. This work has also been extended to generate hybrid semantics in [16, 15, 8]. See also the chapter of this book on CARMA [42].

## 6 Conclusion

In this document, we reviewed the notion of mean field limits of a stochastic population process and presented two extensions. The classical mean field results show that, under some conditions, a stochastic population process converges to a deterministic system of ODEs as the number of objects of the population grows. We have shown that, by replacing the system of ODEs by either a differential inclusion or a hybrid system, it is possible to enlarge the set models for which a mean field limit exists. We illustrated these notions by using a classical SIR example. The last section gives a few pointers to papers in which these frameworks can be applied or generalised.

## References

- [1] H. Andersson and T. Britton. *Stochastic Epidemic Models and Their Statistical Analysis*. Springer-Verlag, 2000.
- [2] J. Aubin and A. Cellina. *Differential Inclusions*. Springer-Verlag, 1984.
- [3] C. Baier et al. “Model-checking algorithms for continuous-time Markov chains”. In: *Software Engineering, IEEE Transactions on* 29.6 (2003), pp. 524–541. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1205180](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1205180).
- [4] M. Benaim and J.-Y. Le Boudec. “A class of mean field interaction models for computer and communication systems”. In: *Performance Evaluation* 65.11 (2008), pp. 823–838.
- [5] P. Billingsley. *Probability and measure*. English. Hoboken, N.J.: Wiley, 2012. ISBN: 9781118122372 1118122372.
- [6] L. Bortolussi and N. Gast. “Mean Field Approximation of Imprecise Population Processes”. In: *QUANTICOL Technical Report TR-QC-07-2015* (2015).

- [7] L. Bortolussi et al. “Continuous Approximation of Collective Systems Behaviour: a Tutorial”. In: *Performance Evaluation* 70.5 (2013), pp. 317–349. ISSN: 0166-5316. DOI: 10.1016/j.peva.2013.01.001. URL: <http://www.sciencedirect.com/science/article/pii/S0166531613000023>.
- [8] L. Bortolussi. “Hybrid behaviour of Markov population models”. In: *Information and Computation* (2015).
- [9] L. Bortolussi. “Limit Behavior of the Hybrid Approximation of Stochastic Process Algebras”. en. In: *Analytical and Stochastic Modeling Techniques and Applications*. Ed. by K. Al-Begain, D. Fiems, and W. J. Knottenbelt. Lecture Notes in Computer Science 6148. Springer Berlin Heidelberg, 2010, pp. 367–381. ISBN: 978-3-642-13567-5 978-3-642-13568-2. URL: [http://link.springer.com/chapter/10.1007/978-3-642-13568-2\\_26](http://link.springer.com/chapter/10.1007/978-3-642-13568-2_26) (visited on 06/11/2015).
- [10] L. Bortolussi and J. Hillston. “Model checking single agent behaviours by fluid approximation”. In: *Information and Computation* 242 (2015), pp. 183–226. ISSN: 0890-5401. DOI: 10.1016/j.ic.2015.03.002.
- [11] L. Bortolussi and R. Lanciani. “Fluid Model Checking of Timed Properties”. In: *Formal Modeling and Analysis of Timed Systems*. Springer International Publishing, 2015, pp. 172–188.
- [12] L. Bortolussi and R. Lanciani. “Model Checking Markov Population Models by Central Limit Approximation”. In: *Quantitative Evaluation of Systems*. Lecture Notes in Computer Science 8054. Springer Berlin Heidelberg, 2013, pp. 123–138. (Visited on 06/11/2015).
- [13] L. Bortolussi and R. Lanciani. “Stochastic approximation of global reachability probabilities of Markov population models”. In: *Computer Performance Engineering*. Springer International Publishing, 2014, pp. 224–239.
- [14] L. Bortolussi and A. Policriti. “Dynamical Systems and Stochastic Programming: To Ordinary Differential Equations and Back”. In: *Transactions on Computational Systems Biology XI* (2009), pp. 216–267.
- [15] L. Bortolussi and A. Policriti. “(Hybrid) automata and (stochastic) programs: The hybrid automata lattice of a stochastic program”. In: *Journal of Logic and Computation* (2013), exr045.
- [16] L. Bortolussi and A. Policriti. “Hybrid Dynamics of Stochastic Programs”. In: *Theoretical Computer Science* 411.20 (2010), pp. 2052–2077. ISSN: 0304-3975.
- [17] A. Chaintreau, J.-Y. Le Boudec, and N. Ristanovic. “The age of gossip: spatial mean field regime”. In: *Proceedings of the ACM SIGMETRICS*. Vol. 37. 1. ACM, 2009, pp. 109–120.
- [18] F. Ciocchetta and J. Hillston. “Bio-PEPA: A framework for the modelling and analysis of biological systems”. In: *Theoretical Computer Science* 410.33 (2009). 00185, pp. 3065–3084. URL: <http://www.sciencedirect.com/science/article/pii/S0304397509001662> (visited on 11/25/2013).
- [19] A. Crudu et al. “Convergence of stochastic gene networks to hybrid piecewise deterministic processes”. In: *The Annals of Applied Probability* 22.5

- (2012). 00015, pp. 1822–1859. URL: <http://projecteuclid.org/euclid.aoad/1350067987> (visited on 11/05/2013).
- [20] R. Darling, J. R. Norris, et al. “Differential equation approximations for Markov chains”. In: *Probability surveys* 5 (2008), pp. 37–79.
- [21] M. H. A. Davis. *Markov Models and Optimization*. Chapman & Hall, 1993.
- [22] J. Doncel, N. Gast, and B. Gaujal. “Mean-Field Games with Explicit Interactions”. working paper or preprint. Feb. 2016. URL: <https://hal.inria.fr/hal-01277098>.
- [23] R. Durrett. *Essentials of Stochastic Processes*. en. Springer, May 2012. ISBN: 9781461436157.
- [24] C. Fricker and N. Gast. “Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity”. In: *EURO Journal on Transportation and Logistics* (2014), pp. 1–31.
- [25] C. Fricker, N. Gast, and H. Mohamed. “Mean field analysis for inhomogeneous bike sharing systems”. In: *DMTCS Proceedings* 01 (2012), pp. 365–376.
- [26] V. Galpin. “Spatial representations and analysis techniques”. In: *SFM*. 2016.
- [27] V. Galpin, L. Bortolussi, and J. Hillston. “HYPE: Hybrid modelling by composition of flows”. In: *Formal Aspects of Computing* 25.4 (2013), pp. 503–541.
- [28] N. Gast and B. Gaujal. “Markov chains with discontinuous drifts have differential inclusion limits”. In: *Performance Evaluation* 69.12 (2012), pp. 623–642.
- [29] N. Gast and B. Gaujal. “Mean field limit of non-smooth systems and differential inclusions”. In: *ACM SIGMETRICS Performance Evaluation Review* 38.2 (2010), pp. 30–32.
- [30] N. Gast, J.-Y. Le Boudec, and D.-C. Tomozei. “Impact of demand-response on the efficiency and prices in real-time electricity markets”. In: *Proceedings of the 5th international conference on Future energy systems*. ACM, 2014, pp. 171–182.
- [31] N. Gast and B. Van Houdt. “Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms”. In: *ACM SIGMETRICS 2015*. 2015.
- [32] J. Hasenauer et al. “Method of conditional moments (MCM) for the Chemical Master Equation: A unified framework for the method of moments and hybrid stochastic-deterministic models”. en. In: *Journal of Mathematical Biology* (Aug. 2013). ISSN: 0303-6812, 1432-1416. DOI: 10.1007/s00285-013-0711-5. URL: <http://link.springer.com/10.1007/s00285-013-0711-5> (visited on 07/31/2014).
- [33] T. Henzinger, B. Jobstmann, and V. Wolf. “Formalisms for Specifying Markovian Population Models”. In: *International Journal of Foundations of Computer Science* 22.04 (2011), pp. 823–841. URL: <http://www.worldscinet.com/doi/abs/10.1142/S0129054111008441>.

- [34] L. Hu, J.-Y. Le Boudec, and M. Vojnoviae. “Optimal channel choice for collaborative ad-hoc dissemination”. In: *INFOCOM, 2010 Proceedings IEEE*. IEEE. 2010, pp. 1–9.
- [35] M. Huang, R. P. Malhamé, P. E. Caines, et al. “Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.
- [36] J.-P. Katoen, M. Khattri, and I. S. Zapreevt. “A Markov reward model checker”. In: *Quantitative Evaluation of Systems, 2005. Second International Conference on the*. 2005, pp. 243–244. (Visited on 01/18/2014).
- [37] T. Kurtz. “Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes”. In: *Journal of Applied Probability* 7 (1970), pp. 49–58.
- [38] M. Kwiatkowska, G. Norman, and D. Parker. “PRISM 4.0: Verification of probabilistic real-time systems”. In: *Computer Aided Verification*. 2011, pp. 585–591. URL: [http://link.springer.com/chapter/10.1007/978-3-642-22110-1\\_47](http://link.springer.com/chapter/10.1007/978-3-642-22110-1_47) (visited on 01/18/2014).
- [39] M. Krn et al. “Stochasticity in gene expression: from theories to phenotypes”. In: *Nature Reviews Genetics* 6.6 (2005), pp. 451–464. ISSN: 1471-0056, 1471-0064. DOI: 10.1038/nrg1615. URL: <http://www.nature.com/doifinder/10.1038/nrg1615> (visited on 02/09/2016).
- [40] J.-M. Lasry and P.-L. Lions. “Mean field games”. In: *Japanese Journal of Mathematics* 2.1 (2007), pp. 229–260.
- [41] J.-Y. Le Boudec. *Performance Evaluation of Computer and Communication Systems*. EPFL Press, Lausanne, Switzerland, 2010.
- [42] M. Loreti. “Modeling and analysis of collective adaptive systems with CARMA and its tools”. In: *SFM*. 2016.
- [43] M. Mitzenmacher. “The power of two choices in randomized load balancing”. In: *Parallel and Distributed Systems, IEEE Transactions on* 12.10 (2001), pp. 1094–1104.
- [44] J. R Norris. *Markov chains*. English. Cambridge University Press, 1998. ISBN: 978-0-511-81063-3 0-511-81063-6.
- [45] J. Pahle. “Biochemical simulations: stochastic, approximate stochastic and hybrid approaches”. en. In: *Briefings in Bioinformatics* 10.1 (Oct. 2008), pp. 53–64. ISSN: 1467-5463, 1477-4054. DOI: 10.1093/bib/bbn050. URL: <http://bib.oxfordjournals.org/cgi/doi/10.1093/bib/bbn050> (visited on 07/14/2014).
- [46] E. Todorov. “Optimal control theory”. In: *Bayesian brain: probabilistic approaches to neural coding* (2006), pp. 269–298.
- [47] M. Tribastone, S. Gilmore, and J. Hillston. “Scalable differential analysis of process algebra models”. In: *Software Engineering, IEEE Transactions on* 38.1 (2012), pp. 205–219. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5567115](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5567115) (visited on 11/24/2013).

- [48] M. Tschaikowski and M. Tribastone. “Approximate Reduction of Heterogeneous Nonlinear Models with Differential Hulls”. In: *submitted to a journal* (2015).
- [49] J. N. Tsitsiklis, K. Xu, et al. “On the power of (even a little) resource pooling”. In: *Stochastic Systems* 2.1 (2012), pp. 1–66.
- [50] B. Van Houdt. “A Mean Field Model for a Class of Garbage Collection Algorithms in Flash-based Solid State Drives”. In: *Proceedings of the ACM SIGMETRICS*. SIGMETRICS '13. Pittsburgh, PA, USA: ACM, 2013, pp. 191–202. ISBN: 978-1-4503-1900-3. DOI: 10.1145/2465529.2465543. URL: <http://doi.acm.org/10.1145/2465529.2465543>.
- [51] D. S. Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman & Hall, 2006.
- [52] T. Yang, P. G. Mehta, and S. P. Meyn. “A mean-field control-oriented approach to particle filtering”. In: *American Control Conference (ACC), 2011*. IEEE, 2011, pp. 2037–2043.
- [53] L. Ying. “On the Rate of Convergence of Mean-Field Models: Stein’s Method Meets the Perturbation Theory”. In: *arXiv preprint arXiv:1510.00761* (2015).