Mean Field Approximation of Uncertain Stochastic Models

Luca Bortolussi
DMG, University of Trieste, Trieste, Italy.
MOSI, Saarland University, Saarbrücken, Germany.
CNR-ISTI, Pisa, Italy.
luca.bortolussi@gmail.com

Nicolas Gast
Inria
Univ. Grenoble Alpes
Grenoble, France
nicolas.gast@inria.fr

***Abstract*—We consider stochastic models in presence of uncertainty, originating from lack of knowledge of parameters or by unpredictable effects of the environment. We focus on population processes, encompassing a large class of systems, from queueing networks to epidemic spreading. We set up a formal framework for imprecise stochastic processes, where some parameters are allowed to vary in time within a given domain, but with no further constraint. We then consider the limit behaviour of these systems as the population size goes to infinity. We prove that this limit is given by a differential inclusion that can be constructed from the (imprecise) drift. We provide results both for the transient and the steady state behaviour. Finally, we discuss different approaches to compute bounds of the so-obtained differential inclusions, proposing an effective control-theoretic method based on Pontryagin principle for transient bounds. This provides an efficient approach for the analysis and design of large-scale uncertain and imprecise stochastic models. The theoretical results are accompanied by an in-depth analysis of an epidemic model and a queueing network. These examples demonstrate the applicability of the numerical methods and the tightness of the approximation.**

***Keywords*—stochastic models; population; parameter estimation; mean-field approximation; differential inclusions;**

I. Introduction

The complexity of the world prevents us to describe all its aspects with full precision. Lack of knowledge and limited computational and intellectual resources force any description of a complex behaviour to be imprecise and uncertain to a certain degree. This is true also when we try to construct mathematical models of complex systems, in domains as diverse as telecommunication, molecular biology, epidemiology. Probabilistic models are one way of representing uncertainties but they depend on parameters which values are never known precisely, and can possibly be estimated from data only with a certain degree of imprecision. Furthermore, mathematical models of systems cannot provide an explicit description of the environment, which always influences the behaviour in uncontrollable ways.

As a typical example, consider a model of epidemic spreading in a population, for instance a model of disease spreading in humans [1] or of a malware spreading in a computer network [2]. The infection rate is a typical parameter of such models which can hardly be known exactly, and is usually estimated from available data about the initial outbreak of the epidemic [1]. However, statistical estimation can never provide an exact value, as we have at disposal only a finite, and often insufficient, amount of data. Moreover, the infection rate itself may depend on environmental factors, which can change arbitrarily in time. For instance, in cholera spreading [3], the level of rainfall impacts on the diffusion of the bacterium among nearby water reservoirs, leading to infection rates which can vary unpredictably in time.

Another example is provided by sociotechnical systems such as bike sharing systems. In such as system, users can pick up or return bikes at any station of their choice. The arrival rate of users in a bike station cannot be assumed fixed during the day, but it will depend on the current hour. The precise form of such a dependency is unknown, as it will be influenced by several factors like the weather, the status of public transportation, the presence of events in the city. Even if we restrict to a small time frame, like the rush hour in the morning, so that we can assume a constant arrival rate, such a rate cannot be reliably fixed to a precise value, but should rather be assumed to lie in an interval of possible values.

In this paper, we distinguish two models of uncertainties
Imprecision: Some parameters ϑ can depend on features of the environment external to the model (like temperature, atmospheric weather, and so on,). ϑ may be subject to variations during the time horizon T of interest, so that considering it as a fixed parameter is an incorrect assumption that can lead to incorrect results. One way to capture such a variability, without committing to assumptions on the form of dependency of ϑ on the external / environmental factors, is to fix a set Θ of possible values for ϑ and assume that ϑ depends on time t and can take any value of Θ at any time instant, i.e. $\vartheta_t \in \Theta$. Essentially, ϑ_t plays a role similar to inputs in control theory. We call this the *imprecise* scenario.
Uncertainty: In a simpler scenario, a parameter ϑ is assumed fixed, but its precise value is not known precisely. In this case, we assume that $\vartheta \in \Theta$, where Θ is the possible set of values of ϑ, as above. This will be referred to as the *uncertain* scenario.

In this paper, we focus our attention on models of large populations of interacting agents, formalised in terms of continuous time Markov chains (CTMC), introducing a new class of models, called *imprecise population processes*, which are able to capture the intrinsic uncertainty of the world. In these models, we assume that some parameters can vary in an unconstrained way within a certain range. This non-deterministic variation represents all possible ways the external environment can influence the value of these parameters, and is a model of

the imprecise scenario described above. We also distinguish a simpler case, in which parameters are assumed to be fixed, i.e. not influenced by the environment, but unknown. In this case, the class of population models so obtained is considerably simpler, resulting in the so-called uncertain continuous time Markov chains, see [4].

When populations are large, as is often the case in epidemiology, biology, telecommunications, a direct analysis of the stochastic model is unfeasible, even by restricting to a statistical treatment of simulations. A viable alternative is to construct the so-called mean field approximation [5], which provides a deterministic description of the behaviour to which stochastic trajectories converge in the limit of infinite populations.

The complexity of the analysis is further exacerbated by the presence of imprecision or uncertainty. The main contribution of this paper is to provide a characterisation of mean-field limits for imprecise population processes (and a-fortiori for uncertain CTMCs) in terms of differential inclusions (DI, [6]), both for transient and steady-state behaviour (when this is meaningful). We investigate the computational gains obtained in this way. With respect to classic mean field, the presence of imprecision increases the computational cost of analysis also for the mean-field limits, as we have to deal with differential inclusions. In the paper, we also present a novel method to bound the solutions of DI using control theoretic tools, namely the Pontraygin principle [7, Section 3].

These results pave the way to novel and efficient computational methods to analyse, design, and eventually control large scale systems of interacting agents, taking both uncertainty in parameter specification and imprecision due to external effects consistently into account. For instance, we can design patching (or vaccination) strategy to counteract an epidemic which is effective even if the infection rate changes in time in unpredictable ways.

The paper is organised as follows: Section II discusses Imprecise Markov Chains and the evolution of their probability mass in terms of differential inclusions. Section III introduces imprecise population models and proves their mean field behaviour. Section IV discusses computational methods for the analysis of the differential inclusion mean field limits. Section V works out in detail an example about epidemic spreading, while Section VI discusses a generalised processor sharing scenario. The final discussion is in Section VII.

Related work: Following the terminology of [8], the models developed in our paper combine two ways for representing uncertainties: a "stochastic uncertainty", driven by a Markovian behaviour, and a "contingent uncertainty", given by possibly changing parameters. This approach is similar to the notion on stochastic differential inclusion studied in [8], [9], where the "stochastic uncertainty" is driven by a Wiener process while the "contingent uncertainty" is given by a set-valued map. Our work also builds on [10], that considers Markov chain with interval probabilities. The key contribution of our paper with respect to [10] is to extend this notions to population process and to develop a rigorous mean-field approximation of such systems.

Mean field approximation of population processes has a long history, starting from the works of Kurtz [11], [12]. Mean field based analysis, for the transient and the steady state, have been applied in performance modelling and model-checking tools [13], [14], systems biology, epidemiology [5]. For a gentle introduction, see [15].

Classic mean field results require (Lipschitz) continuity of rate functions, but more general theorems can be proved for piecewise smooth rates [16] or even general discontinuous functions by using differential inclusions [17]. The proof of most of our convergence results are based on the construction of a proper stochastic approximation. They are similar to the ones of [17], [18]. For clarity and brevity, we will only give sketches of proofs.

Note that the main technical contribution of this paper is to properly define the mean-field framework for uncertainties, in a way that makes feasible the adaptation of proofs from stochastic approximation by differential inclusions [18].

II. Imprecise Markov Chains

In this section we first discuss imprecise continuous-time Markov chains in general. Population processes are a subclass of this general model. We will introduce the imprecise and the uncertain models. We will then introduce briefly differential inclusions, and show how Kolmogorov equations for the probability mass generalise to differential inclusions in this setting. We also introduce the imprecise drift of an imprecise model, which will play a central role in the construction of mean field limits.

A. Imprecise and uncertain Markov chains

We consider a stochastic process $\mathbf{X} = (X_t)_{t\geq 0}$ that takes value in a *discrete* state space $\mathbf{E} \subset \mathbb{R}^d$ and that is adapted to a filtration[1] $\mathcal{F}$. The dynamics of the process depends on a parameter (or a vector of parameters) ϑ. We denote by Θ the set of possible parameter values of ϑ. We consider a set of transitions kernel Q^ϑ on $\mathbf{E}$, parametrized by $\vartheta \in \Theta$: For each $\vartheta \in \Theta$, Q^ϑ is such that $Q^\vartheta_{x,y} \geq 0$ for $x \neq y \in \mathbf{E}$ and $\sum_{y\in\mathbf{E}} Q^\vartheta_{xy} = 0$.

Definition 1. *An imprecise continuous time Markov chain is a stochastic process $\mathbf{X}$ together with a $\mathcal{F}_t$-adapted process θ such that for all $t \geq 0$:*

$$\lim_{h\to 0} \frac{1}{h}\mathbb{P}(X_{t+h} = y \mid \mathcal{F}_t, X_t = x) = \begin{cases} Q^{\vartheta_t}_{xy} & \text{if } x \neq y \\ -\sum_{y\neq x} Q^{\vartheta_t}_{xy} & \text{if } x = y \end{cases}$$

The definition of an imprecise Markov chain makes no restriction on the set of processes that the varying parameter θ can take. In some cases, it can be interesting to focus on subset of processes. An example is if we assume that ϑ_t is deterministic and constant in time. In that case we obtain the notion of uncertain Markov chain:

Definition 2. *An uncertain continuous time Markov chain is a stochastic process $\mathbf{X}$ such that there exists a constant parameter $\vartheta \in \Theta$ such that for all $t \geq 0$:*

$$\lim_{h\to 0} \frac{1}{h}\mathbb{P}(X_{t+h} = y \mid \mathcal{F}_t, X_t = x) = \begin{cases} Q^{\vartheta}_{xy} & \text{if } x \neq y \\ -\sum_{y\neq x} Q^{\vartheta}_{xy} & \text{if } x = y \end{cases}$$

[1] A filtration is a set of σ-algebra $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for each $t \geq s \geq 0$. $\mathbf{X}$ is adapted to $\mathcal{F}$ means that X_t is $\mathcal{F}_t$ measurable for every t.

The definitions of imprecise and uncertain CTMC correspond to two extreme cases: in the imprecise case, the parameter function θ can be any measurable function while in the uncertain case the parameter function θ is deterministic and constant in time. It would be possible to consider cases in between, for example by restricting the set of admissible processes θ to be the set of functions that only depends on the value $X(t)$ – such process θ are called Markovian control policies in the Markov decision processes community – or the set of time-dependent deterministic functions – which would lead to time-inhomogeneous CTMC.

Remark: Imprecise CTMC are strongly related to (continuous-time) Markov Decision Processes (MDPs) [19], the main difference being that in MDPs, the emphasis is on finding a policy that maximises some reward criteria. To to this, one usually assumes a finite or countable number of actions that can be taken non-deterministically. In ICTMC, the variability of the decision variables is usually not controllable and can be a generic random function adapted to the process. The parameter space is in general uncountable and our objective is to characterise the set of possible behaviours. Uncertainty, in the sense considered here, has been considered also in the MDS context, in the so called Uncertain (Bounded-Parameter) MDPs [20], [21].

Example: For illustrative purposes, we consider a simple model of a bike sharing system, similar to the one of [22]. We describe the number of bikes present in a single station, so that the state of the CTMC is given by $X_t \in \{0, \dots, N\}$, where N is the capacity of the station, i.e. the maximum number of bikes. We assume that customers arrive at an unknown rate ϑ_a, belonging to the interval $[\vartheta_a^{min}, \vartheta_a^{max}]$. Each customer arrival brings the system from state k to $k-1$ (for $k > 0$). Similarly, we can model the return of a bike as a transition with rate ϑ_r, belonging to the interval $[\vartheta_r^{min}, \vartheta_r^{max}]$, and increasing the number of available bikes from k to $k+1$, provided $k < N$. Different choices of how the two imprecise parameters can vary give rise to different models. In the general case, we can assume that $\vartheta_a(t)$ and $\vartheta_r(t)$ are generic functions of time, encoding complex but unknown dependencies of the environment of customers requesting a bike and wishing to return it at the station. On the opposite spectrum, we could assume that these values are unknown but constant in time, for example if we are considering the dynamics of the station in a restricted time frame.

B. Differential inclusions

Differential inclusions (DI) are a generalisation of differential equations, and provide the natural mathematical tool to describe the transient evolution of the probability mass of an imprecise CTMC. Furthermore, as we will see in the next section, under a proper scaling, the behaviour of an imprecise Markov chain is closely related to the one of a differential inclusion corresponding to Equation (3). In this section, we provide a quick introduction to DI, recalling some classical definitions. See [6] for further details.

Let F be a set-valued function on $E \subset \mathbb{R}^d$ that assigns to each $x \in E$ a set of vectors $F(x) \subset \mathbb{R}^d$. A solution to the differential inclusion $\dot{x} \in F(x)$ that starts in x is a function $\mathbf{x} : [0, \infty) \to E$ such that there exists a measurable function $f : [0, \infty) \to \mathbb{R}^d$ satisfying: for all $t \geq 0$ $f(t) \in F(x(t))$ and

$$x(t) = x_0 + \int_0^t f(s)ds.$$

For an initial condition x_0, we denote by S_{F,x_0} (or S_{x_0} if there is no ambiguity) the set of solutions of $\dot{x} \in F(x)$ that start in x_0. Note that the set S_{F,x_0} can be empty or be composed of multiple solutions, depending on the function F. When $E = \mathbb{R}^d$, a sufficient condition for the existence of at least one solution is that (a) for all $x \in \mathbb{R}^d$ $F(x)$ is non-empty, convex and bounded (*i.e.*, $\sup_{x \in X, y \in F(x)} \|y\| < \infty$) and (b) F is upper-semi-continuous (*i.e.*, the graph of F, $\{(x, y), y \in F(x)\}$, is a closed set).

Asymptotic behaviour as t goes to infinity: As the time grows, a solution $\mathbf{x}$ of a differential inclusion can have a chaotic behaviour and can oscillate around many points. For a given starting point x_0, we define the limit set $L_F(x_0)$ as the closure of the set of points that are accumulation points of at least one solution of the differential inclusion starting in x_0:

$$L_F(x_0) := \bigcup_{\mathbf{x} \in S_{F,x_0}} \bigcap_{t \geq 0} \text{closure}(\{x(s), s \geq t\}),$$

where *closure* denotes the closure in the metric space $\mathbb{R}^d$.

As in [18], we call the Birkhoff centre of the differential inclusion is the closure of the set of recurrent points of the differential inclusion:

$$B_F = \text{closure}(\{x \in E : x \in L_F(x)\}). \tag{1}$$

Intuitively, the Birkhoff centre contains all attractors, equilibria, limit cycles, and in general all points that can be visited infinitely often by one solution of the differential inclusion. As such, it provides a characterisation of the localisation in the phase space of the limit behaviour of the system. Note that, in general, the sets $L_F(x)$ and B_F are not convex nor necessarily connected, even when $F(x)$ is single valued and Lipschitz (for example the Birkhoff centre of $F(x) = \{x(1-x)\}$ is $\{0, 1\}$). Note that even if the case of Lipschitz ordinary differential equation, the computation of the Birkhoff centre of an ODE is complicated and even its shape can depend strongly on the parameter of the function (see the comparison between Figure 1 and Figure 2 of [5]).

C. Imprecise Kolmogorov Equations

The time evolution of the probability mass of an imprecise Markov chain can be obtained by a generalised version of Kolmogorov equations, stated in terms of differential inclusions. For a fixed realisation ϑ_t of the process ϑ, the evolution of the probability mass $P(t \mid \vartheta_t)$ is given by the standard (non-autonomous) Kolmogorov (forward) equations $\dot{P}(t \mid \vartheta_t) = Q^{\vartheta_t} \cdot P(t \mid \vartheta_t)$. However, if we have no knowledge on the process ϑ, we can only assume that the process at time t behaves like a Q^{ϑ} for some ϑ, implying that the transient behaviour of the probability mass will be a solution of the differential inclusion

$$\dot{P}(t) \in Q \cdot P(t), \tag{2}$$

where $Q = \bigcup_{\vartheta \in \Theta} Q^{\vartheta}$. Equation (2) is linear, but the dimension of the differential inclusion equals the size of the state space, which is too large in most interesting scenarios, especially for population models, see Section III.

D. Generalised Drift

An important notion associated with an imprecise Markov chain is that of drift, which extends the corresponding notion for CTMCs, and describes the average increment in a point of the state space, as a function of the imprecise parameters.

Definition 3. *The imprecise drift of the imprecise CTMC is the function $f : E \times \Theta \to \mathbb{R}^d$, defined by*

$$f(x, \vartheta) = \sum_{y \in E} Q^{\vartheta}_{x,y}(y - x). \tag{3}$$

When for all $x \in E, \vartheta \in \Theta$ we have $\sum_{y \in E} Q^{\vartheta}_{x,y} \|y - x\| < \infty$, the sum in Equation (3) *is well defined.*

III. Imprecise Population Processes

In this section we introduce the class of models of interest in this paper, namely imprecise population processes. Population processes are a very common class of models, and when relaxing the precision intrinsic in their definition, Imprecise population processes emerge naturally. After providing the definition, we shift the focus to the behaviour of imprecise population processes in the limit of an infinite population. We show that such a limit is described by a differential inclusion, which provides also information about the stationary behaviour of the imprecise process, encoded in the Birkhoff centre.

A. Definition

Let N be a scaling parameter (typically, N is the population size of the considered model). We consider a sequence of imprecise Markov processes indexed by N, denoted $(\mathbf{X}^N)_{N \geq 0}$ on a sequence of subset $\mathbf{E}^N \subset E \subset \mathbb{R}^d$. The stochastic process $\mathbf{X}^N$ is an imprecise process of kernel $Q^{N,\theta}$ and $X^N(t)$ denotes the state of the process at time t.

Definition 4. *An imprecise (respectively uncertain) population process is a sequence of imprecise (respectively uncertain) Markov chains that satisfies the following assumptions:*

(i) *The chains are uniformizable:* i.e.*, for all N, the rates are uniformly bounded:* $\sup_{x \in E^N, \vartheta \in \Theta} |Q^{N,\vartheta}_{xx}| < \infty$

(ii) *The transitions become smaller as N grows,* i.e.*, there exists $\varepsilon > 0$ such that*

$$\lim_{N \to \infty} \sup_{x \in E^N, \vartheta \in \Theta} \sum_{y \in E^N} Q^{N,\vartheta}_{xy} \|y - x\|^{1+\varepsilon} = 0$$

(iii) *The drifts are well-defined and bounded:*

$$\limsup_{N \to \infty} \sup_{x \in E^N, \vartheta \in \Theta} \sum_{y \in E^N} Q^{N,\vartheta}_{xy} \|y - x\| < \infty$$

Transition classes: A direct definition of an imprecise population process by exhibiting its generator is unfeasible. A simpler definition can be obtained by specifying transition classes, similarly to [15], and describing the state of the system by a state vector, counting the size of each population. The idea is to specify the dynamics by a set of possible transitions or events, providing their rate, as a function of the state vector and of the imprecise parameters, and how they change the state vector of the system. Specification can be done at the level of the density (normalised population process) or on the integer-valued counting variables (non-normalised process). Usually, such specifications satisfy the conditions in the previous definition. See below and Section V for examples.

Example: A simple example of a population model is provided by a slight modification of the single station bike sharing system discussed in the previous section. In this case, we can assume the total population N is given by the total amount of bike racks in the station, while the only model variable, $X_B(t)$, encodes the number of bikes available at the station in terms of the fraction of occupied bike racks. There are two transitions classes in the model: a customer arrival, taking one bike, and a biker arrival, returning one bike. The former transition class has rate $N\vartheta_a(t)$, if $X_B(t) > 0$, and changes the state of the system from X_B to $X_B - 1/N$. Similarly, bike arrivals happens at rate $N\vartheta_r(t)$, if $X_B(t) < 1$, and move the system from X_B to $X_B + 1/N$. The dependency of the arrival rates on N is needed to satisfy the conditions of the definition above, and intuitively describes the idea that ϑ_a and ϑ_r model the relative traffic volume in terms of bike density, so that the absolute traffic level is obtained by multiplication with the system size N.

B. Mean field limit (finite time-horizon)

Let $f^N(x, \vartheta) = \sum_{y \in E^N} Q^{N,\vartheta}_{xy}(y - x)$ be the drift of the system N. We define the limit drift of the system as the convex closure of the set of the accumulation points of $f^N(x^N, \vartheta)$ as N goes to infinity, for all sequences x^N that converge to x:

$$\begin{aligned} F(x) &= \overline{\lim_{N \to \infty}} \bigcup_{\vartheta \in \Theta} \left\{ f^N(x, \vartheta) \right\}, \\ &= \Bigg\{ y \in \mathbb{R}^d \text{ such that there exists a sequence } \phi_N : \\ &\qquad \lim_{N \to \infty} \phi_N = \infty \wedge \lim_{N \to \infty} f^{\phi_N}(x^{\phi_N}, \vartheta^{\phi_N}) = y \Bigg\}. \end{aligned} \tag{4}$$

As in [17], it can be shown that when the $Q^{N,\vartheta}$'s satisfy Definition 4, the differential inclusion $\dot{x} \in F(x)$ has a unique solution as for all $x \in \mathbb{R}^d$ $F(x)$ is non-empty, convex and bounded (*i.e.*, $\sup_{x \in X, y \in F(x)} \|y\| < \infty$) and F is upper-semi-continuous.

Theorem 1. *Let $(\mathbf{X}^N)$ be an imprecise population process. Then, if $X^N(0)$ converges (in probability) to a point x, then the stochastic process $\mathbf{X}^N$ converges (in probability) to $S_{F,x}$, the set of solutions of $\dot{x} \in F(x)$ starting in x.*

Sketch of proof: The difficult part of the proof is to show that there exists a decreasing function $I(N)$ such that the discrete time process $\mathbf{X}^N(kI(N))$ satisfies the definition of a GASP of [18]. This can be done similarly to [17]. Hence, the process $(\mathbf{X}^N(kI(N)))_k$ satisfies the assumptions of [18, Theorem 3.1] which concludes the proof. ∎

Remark: The construction of the imprecise drift generalises the definition given in [17], where the focus was on discontinuous, but precise, rates. The proof of the theorem follows the ideas of [17], [23]. In fact, when Θ is restricted to a single value, we obtain again the result of [17].

In the case of an uncertain population process, we define the drift $F_\vartheta(x)$ as a function of the state $x \in E$ and the

parameter $\vartheta \in \Theta$.

$$F_\vartheta(x) = \overline{\lim_{N\to\infty}} \left\{ f^N(x,\vartheta) \right\}. \tag{5}$$

It should be clear that for all x: $\bigcup_{\vartheta\in\Theta} F_\vartheta(x) \subset F(x)$. Note that the quantity $F_\vartheta(x)$ is non-necessarily single-valued when the dynamics is not-continuous (see [17]).

Corollary 1. *Let $\mathbf{X}^N$ be an uncertain Markov population process, then if $X^N(0)$ converges (in probability) to a point x, then the stochastic process $\mathbf{X}^N$ converges (in probability) to $\bigcup_\vartheta S_{\vartheta,x}$, where $S_{\vartheta,x}$ is the set of solutions of the differential inclusion $\dot{x} \in F_\vartheta(x)$ starting in x.*

These results are illustrated in Section V.

C. Mean field limit (stationary regime)

An uncertain Markov chain does not necessarily have a stationary behaviour. In particular, an imprecise Markov chain is not necessarily time-homogeneous (the parameter ϑ_t can depend on time) and is also not necessarily a Markov chain (ϑ_t can depend on all the past of the stochastic process). Nevertheless, it is possible to constraint the asymptotic regime of an imprecise Markov population process by using the asymptotic reachable set of the differential inclusion, A_F, defined by:

$$A_F = \bigcap_{T>0} \bigcup_{x_0, t\geq T, \mathbf{x}\in S_{F,x_0}} \{\mathbf{x}(t)\} \tag{6}$$

Note that the the Birkhoff centre B_F is included in the set A_F. The inclusion is in general strict.

The following result is a direct consequence of Theorem 1.

Theorem 2. *Let $(\mathbf{X}^N)_N$ be an imprecise population process, then there exists a sequence ε^N that converges to 0 in probability and such that the distance between $X^N(t)$ and A_F becomes stochastic lower than ε^N as t goes to infinity. In other word:*

$$\lim_{N\to\infty} \lim_{t\to\infty} d(X^N(t), A_F) = 0 \qquad \textit{in probability}.$$

When in addition of being an imprecise Markov chain, for all N, the process $\mathbf{X}^N$ is a Markov chain and has a stationary measure μ^N, we can say more. The following results characterise the stationary behaviour of the stationary measures as N grows. It shows that the sequence of stationary measures μ^N concentrates on the Birkhoff centre of the differential inclusion. The Birkhoff centre, defined in Section II-B, is essentially the set of recurrent points of the differential inclusion, *i.e.*, the set of points such that there exists a trajectory that starts at this point and comes back to this point in the future.

Theorem 3. *Let $\mathbf{X}$ be an imprecise population process such that $\mathbf{X}^N$ is a Markov chain that has a stationary measure μ^N. Let μ be a limit point of μ^N (for the weak convergence). Then, the support of μ is included in the Birkhoff centre of F, defined in Equation* (1)*: $\mu(B_F) = 1$.*

Sketch of proof: As for the proof of Theorem 1, one can show that there exists a function $I(N)$ such that $\mathbf{X}^N(kI(N))$ is a GASP. Theorem 3 is therefore a consequence of [18, Corollary 9]. ∎

Corollary 2. *Let $\mathbf{X}$ be an uncertain population process such that $\mathbf{X}^N$ is a Markov chain that has a stationary measure μ^N. Let μ be a limit point of μ^N (for the weak convergence). Then, the support of μ is included in the Birkhoff centre of F_ϑ, defined in Equation* (1)*.*

Remark: Theorem 3 only states that the support of μ is included in the Birkhoff centre but provides no intuition on the how the probability mass is spread on this set. This results can be refined by using the notion of semi-invariant measure [18, Definition 3.3]. However, this notion is complex and computing a semi-invariant measure of a differential inclusion seems numerically intractable, as it requires to compute a probability measure on all the possible trajectories of the differential inclusion. Hence, it the present document, we limit our exposition to the notion of Birkhoff centre, which, even if less accurate, provides a simpler characterisation that can be computed numerically.

IV. Numerical Methods

The results of the previous section imply that, for large populations, we can study the mean field differential inclusion to get insights on the transient and on the steady state of the stochastic population process. The analysis of this class of mean field models, however, is in general more challenging than ordinary differential equations (ODE). In the differential inclusion case, in fact, we usually are only able to compute bounds on the solution set of the equations. After discussing some existing approaches, we present in more detail two fast methods, one based on differential hulls and the other based on the control-theoretic Pontraygin principle.

A. Related work on numerical methods

Most recent numerical approaches dealing with imprecise deterministic processes have been developed for computing reachable sets of hybrid systems, whose continuous dynamics can be specified by (non-linear) differential equations or differential inclusions. The proposed methods in literature can be roughly divided in two classes: exact over-approximation methods and simulation-based methods [24].

The first class of methods manipulates directly sets of states, finitely represented, for instance, as polytopes [25], ellipsoids [26], or zonotopes [27], or by relying on interval arithmetic [28] combined with constraint solving [29], [30], exploiting satisfaction modulo solvers over reals [31], [32]. The dynamics of the system is lifted at the set level, so that one computes the evolution of the reachable set under the action of the dynamics. These methods usually compute over-approximation of the real reachable set, which is formally guaranteed to contain all reachable points. Most of the methods in literature are restricted to linear systems, due to their wide diffusion in engineering applications. For this class of systems, efficient methods exist [27], [33]. Tools like SpaceEx [34] implement routines for linear differential inclusions capable of solving problems up to several hundred dimensions. Over-approximation methods for non-linear systems, instead, are much less developed, due to the intrinsic difficulty of the problem. Here we recall hybridization [35], which is based on a localised linearisation of the dynamics. Another family of methods, very common for differential inclusions, is based

on interval-based methods combined with constrain solving to control the over-approximation introduced by interval computations. Among them, we recall [29] and [30]. Other methods for integrating differential inclusions rely on error bounding [36].

An alternative approach is offered by simulation methods, which try to infer the reachable set from few simulated trajectories of the system. These methods are very effective for the limit dynamics emerging from uncertain population processes. Here we recall [37], which uses the sensitivity of the system with respect to initial conditions to (approximately) compute how a set of initial conditions propagates in T units of time under the action of the dynamics. The approach of [38], instead, use statistical methods borrowed from machine learning to perform inference of the reachable set with statistical error guarantees. Other methods work for more general imprecise limit models; for example, the procedure of [39]–[41] constructs an under-approximation of the reachable set using a Monte-Carlo sampling method.

B. Differential hull of a differential inclusion

A first method to compute the set of reachable points by the solutions of a differential inclusion is to construct a rectangular over-approximation, which is composed of two functions $\underline{x}$ and $\overline{x}$ such that $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$, for any x that is solution of the differential inclusion x. In this section, we show that we can construct two differential equations such that $\underline{x}$ and $\overline{x}$ are their solutions and such that for all t and any solution $\mathbf{x}$ of the differential inclusion $\dot{x} \in F(x)$, we have $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ (coordinate-wise). The construction of these equations is simple. This bounds are reasonably tight when the set of possible parameters θ is small. However, as we will see in Section V-D, the bounds provided by this approximation become loose when the set of possible ϑ increase.

Following the definition of [13], we say that a locally Lipschitz-continuous function $(\underline{f}, \overline{f})$ is a differential hull for the differential inclusion F if for each coordinate i and each x such that $\underline{x} \leq x \leq \overline{x}$, we have:

$$x_i = \underline{x}_i \Rightarrow \underline{f}_i(\underline{x}, \overline{x}) \leq \inf_{f \in F(x)} f_i$$
$$x_i = \overline{x}_i \Rightarrow \overline{f}_i(\underline{x}, \overline{x}) \geq \sup_{f \in F(x)} f_i.$$

The tightest functions $\underline{f}$ and $\overline{f}$ that are a differential hull for a differential inclusion F are:

$$\underline{f}_i(\underline{x}, \overline{x}) = \min_{x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t)} \min F_i(x)$$
$$\overline{f}_i(\underline{x}, \overline{x}) = \max_{x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t)} \max F_i(x)$$

This leads to the following result, that is a consequence of [42, Theorem 1].

Theorem 4. *Let $\mathbf{x} : [0 : T] \to \mathbb{R}^d$ be a solution of $\dot{x} \in F(x)$ with initial condition $x(0) = x_0$. Let $(\underline{x}, \overline{x})$ be the solution of the differential equation $\dot{\underline{x}} = \underline{f}(\underline{x}, \overline{x})$ and $\dot{\overline{x}} = \overline{f}(\underline{x}, \overline{x})$ with initial condition $\underline{x}(0) = \overline{x}(0) = x_0$. Then, for all t, we have:*

$$\underline{x}(t) \leq x(t) \leq \overline{x}(t)$$

C. Reachability as an optimal control problem: an algorithm based on Pontryagin's maximum principle

In general, the bounds $\underline{x}$ and $\overline{x}$ found by the differential hull method are not tight. We can do better by exploiting Pontryagin's maximum principle, a classical method that can be used to compute numerically the exact minimal value $x^{\min}(t)$ and maximal value $x^{\max}(t)$ that can be reached by an imprecise fluid model at time t [7].

Let $T \geq 0$ be some fixed time and $i \in \{1 \dots d\}$ a coordinate. Let $x_i^{\max}(T) = \sup_{\mathbf{x} \in S_{F,x_0}} x_i(T)$ be the maximal value that the ith coordinate of the solution of a differential inclusion can take at time t. The quantity $x_i^{\max}(T)$ is the solution of the maximization problem over functions $\vartheta(t)$:

$$x_i^{\max}(T) := \max_{\theta} x_i(T) \text{ such that}$$
$$\text{for all } t \in [0;T]: \begin{cases} x(t) = x + \int_0^t f(x(s), \theta(s))ds \\ \theta(t) \in [\vartheta_{\min}, \vartheta_{\max}] \end{cases}$$

Pontryagin's maximum principle is a set of necessary conditions that the trajectory that attains the maximum should satisfy. Following the description of [7, Section 3] these conditions are the following. If x is a trajectory that maximises $x_i(T)$, then there exists a *costate trajectory* $\mathbf{p}$ such that $p_i(T) = -1$, $p_j(T) = 0$ for $j \neq i$ and:

$$\dot{x}(t) = f(x(t), u(t)) \tag{7}$$
$$\vartheta(t) \in \arg\max_{\vartheta} f(x, \vartheta)^T p, \tag{8}$$
$$-\dot{p}(t) = \frac{\partial}{\partial x} \left(f(x, \vartheta)^T p \right) \tag{9}$$

where $f(x, \vartheta)^T p$ denotes the scalar product between $f(x, \vartheta)$ and p.

This formulation can be used to derive a fixpoint algorithm to iteratively obtain x and ϑ. We start from an (arbitrary) initial costate p, satisfying the constraints at time T, then alternatively compute a trajectory x forward in time the x by using (7) and (8), holding p fixed, and computing a new p backward in time by solving (9), holding x fixed. These two steps are repeated until a convergence criterion is met (typically, the sup norm of p and x varies less than a prescribed ε after an iteration).

Remark: The rectangle delimited by $x^{\min}(T)$ and $x^{\max}(T)$ provides an approximation of the reachable set of $x(T)$. This set is tighter than the rectangle delimited by $\underline{x}$ and $\overline{x}$, but is not exact: in general, the set of reachable points is not a rectangle. However, the algorithm given by the iterations (7-9) can be easily extended to refine the rectangle into any convex template polyhedron by considering the minimisation problems $\min_i \sum_i \alpha_i x_i(t)$ for any tuple of coefficients α_i.

V. Illustrative Example 1: the SIR model

In this section, we apply our techniques to the well-known susceptible-infected-recovered (SIR) model. This example will serve us to illustrate the differences between the imprecise and uncertain model and to evaluate our numerical algorithms. We show in particular that the imprecise model can be solved numerically and that it provides bounds that are much tighter than the differential hull approximation.

A. The SIR model

We consider a system composed of N nodes. Each node is in one of the three states: susceptible, infected or recovered. We denote by $X_S(t) \in [0,1]$ the proportion of nodes that are susceptible at time t (and by $X_I(t)$ and $X_R(t)$ the proportion of infected and recovered nodes). By definition, for each t, $X_S(t) + X_I(t) + X_R(t) = 1$.

We model the dynamics of the stochastic system as follows. A susceptible node can become infected from an external source (this occurs at rate a). An infected node becomes recovered at rate b and a recovered node becomes susceptible at rate c. We consider that susceptible and infected node are moving but we do not know at which speed. We model this by considering that a susceptible node encounters an infected node and becomes infected at rate $\vartheta X_I(t)$.

The transitions of the system are the following: The state X_S, X_I, X_R becomes:

- $X_S - \frac{1}{N}, I + \frac{1}{N}, X_R$ at rate $N(aX_S + \vartheta X_S X_I)$,
- $X_S, X_I - \frac{1}{N}, X_R + \frac{1}{N}$ at rate NbX_I and
- $X_S + \frac{1}{N}, X_I, X_R - \frac{1}{N}$ at rate NcX_R.

The sequence of processes $(X_S^N, X_I^N, X_R^N)_N$ satisfies Definition 4 and is therefore an imprecise population process. Its drift is given by the triple $(f_S(X_S, X_I, X_R; \vartheta), f_I(X_S, X_I, X_R; \vartheta), f_R(X_S, X_I, X_R; \vartheta))$, where:

$$\begin{aligned} f_S(X_S, X_I, X_R; \vartheta) &= -aX_S - \vartheta X_S X_I + cX_R \\ f_I(X_S, X_I, X_R; \vartheta) &= aX_S + \vartheta X_S X_I - bX_I \\ f_R(X_S, X_I, X_R; \vartheta) &= bX_I - cX_R \end{aligned} \tag{10}$$

As $X_S + X_I + X_R = 1$, we can substitute for X_R and express the drift as

$$\begin{aligned} f_S(X_S, X_I; \vartheta) &= c - (a+c)X_S - cX_I - \vartheta X_S X_I \\ f_I(X_S, X_I; \vartheta) &= aX_S + \vartheta X_S X_I - bX_I \end{aligned} \tag{11}$$

By Theorem 1 and Corollary 1, as the number of object goes to infinity, the behaviour of the uncertain model (an unknown but constant ϑ) converges to the ODE

$$\dot{x} = f(x, \vartheta)$$

while the imprecise model (a unknown θ that can vary in time) is included in the solution of the differential inclusion

$$\dot{x} \in \bigcup_{\vartheta \in [\vartheta_{\min}, \vartheta_{\max}]} \{f(x, \vartheta)\}$$

In what follows, we will compare numerically the behaviour of the two models. In particular, we show that despite the fact that the infection rate minus the recovery rate $f_I(X_S, X_I; \vartheta) = aX_S + \vartheta X_S X_I - bX_I$ is an increasing function of ϑ, the quantity $x_I(t)$ is not a monotone function of ϑ. In particular, when $\vartheta(t) \in [\vartheta \min, \vartheta \max]$, the proportion of infected nodes can be higher for an imprecise population process than for any uncertain population process.

In the rest of this section, we set the parameters equal to $a = 0.1, b = 5, c = 1, \vartheta_{\min} = 1, \vartheta_{\max} = 10$ and the initial conditions are $X_S(0) = 0.7, X_I(0) = 0.3, X_R(0) = 0$.

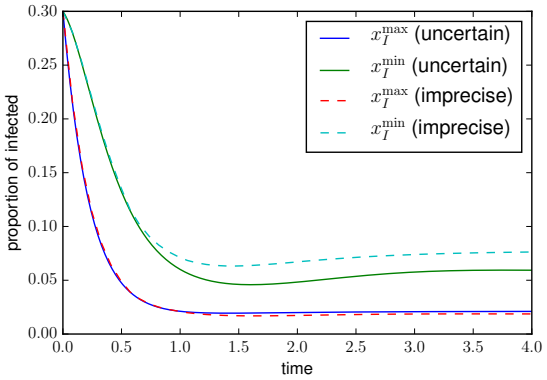


Fig. 1: Upper and lower bounds on the number of infected nodes for an imprecise model (dashed lines) and an uncertain model (solid lines)

B. Reachable sets in finite time

For a fixed parameter ϑ, the drift $f(X_S, X_I, \vartheta)$ is Lipschitz-continuous in X_S and X_I. Hence, the drift of the uncertain model, defined in Equation (5), is a single-valued function and the ODE $(\dot{S}, \dot{I}) = f(X_S, X_I, \vartheta)$ has a unique solution, which we denote by $X_S^\vartheta, X_I^\vartheta$. This ODE does not have a close-form solution but numerical integration is easy.

The case of the imprecise model is more complicated. By Theorem 1, whatever are the variation of the parameter θ (even if it ϑ_t depends on the whole history of the process X_S^N, X_I^N), the quantities X_S^N and X_I^N converges to a solution of the differential inclusion $(\dot{x_S}, \dot{x_I}) \in \{f(x_S, x_I, \vartheta) : \vartheta \in [\vartheta_{\min}, \vartheta_{\max}]\}$. In particular, if S_F is the set of solutions of the differential inclusions and $R_F(t)$ is the set of reachable points at some time t by the differential inclusion starting in $(X_S(0), X_I(0))$, we have:

$$\lim_{N \to \infty} (X_S^N(t), X_I^N(t)) \in R_F(t) := \bigcup_{(s,i) \in S_F} \{(s(t), i(t))\}$$

It should be clear that the set of the possible values for (X_S, X_I) of an uncertain model is included in the set of reachable points of the imprecise model:

$$\bigcup_{\vartheta \in [\vartheta_{\min}, \vartheta_{\max}]} \left\{(x_S^\vartheta(t), x_I^\vartheta(t))\right\} \subset R_F(t). \tag{12}$$

As we show numerically in our example (see Figure 1), the inclusion is, in general, strict.

For both models (uncertain and imprecise), we define the maximum proportion of infected nodes by

$$\begin{aligned} \overline{x}_I^{uncertain}(t) &= \max_\vartheta I_\vartheta(t) \\ \overline{x}_I^{imprecise}(t) &= \max_{(X_S, X_I) \in S_F} X_I(t). \end{aligned}$$

The definition of the minimum proportion of infected node is similar.

The computation of $\overline{x}_I^{uncertain}(t)$ can be done by a numerical exploration of all the parameters ϑ, or relying on the statistical method of [38]. For the computation $\overline{x}_I^{imprecise}(t)$,

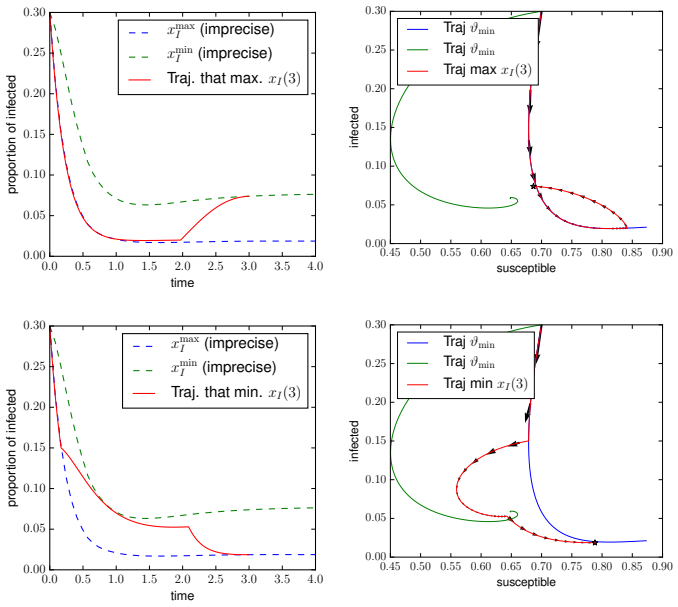


Fig. 2: Trajectory that reaches the maximum (top) or minimum (bottom) number of infected node at time $T = 3$. On the left, we plot the number of infected as a function of time. On the right, we plot the number of infected as a function of the number of susceptible.

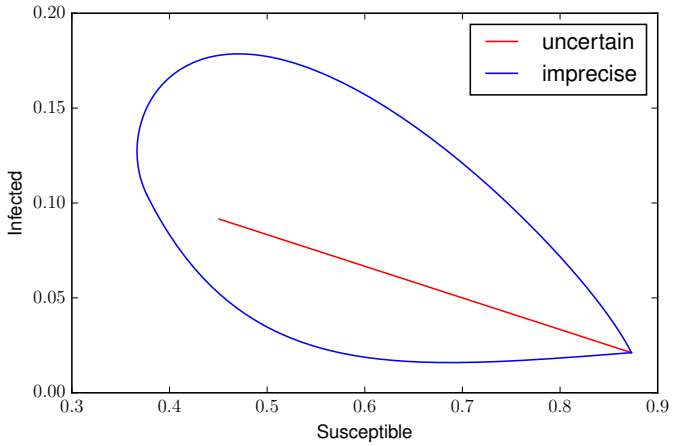


Fig. 3: Steady-state regime for the imprecise and uncertain SIR models. The steady-state of the imprecise model is the convex set delimited by the blue region. The steady-state of the uncertain model is on the red line. We set $\vartheta_{\max} = 10\vartheta_{\min}$.

we formulate the problem as an optimal control problem which is to find a function $\theta(t)$ that maximises $X_I(t)$, where $X_I(t)$ satisfies $\dot{x_S}, \dot{x_I} = f(x_S(t), x_I(t), \theta(t))$. We use a numerical method based on Pontryagin's maximum principle (see Section IV-C). The results are reported in Figure 1. We observe that $\overline{x}_I^{imprecise}(t)$ can be much larger than $\overline{x}_I^{uncertain}(t)$, especially for large value of t.

On Figure 2, we add to Figure 1 two examples of trajectories that attain the maximum (top) or minimum (bottom) number of infected nodes at time $T = 3$. These maximal and minimal trajectories are in fact solutions of bang-bang control policies. For the trajectory that maximises $X_I(3)$, the control $\theta(t)$ equals $\vartheta_{\min}$ for $t < 2.25$ and $\vartheta_{\max}$ for $t > 2.25$. For the trajectory that minimises $X_I(3)$, the control $\theta(t)$ is to start with $\vartheta_{\min}$ for $t < 0.7$, then use $\vartheta_{\max}$ until $t = 2.2$ and then use again $\vartheta_{\min}$.

C. Steady-state regime

Computing the Birkhoff centre B_F of a differential inclusion is in general not easy. A first possibility is to use Pontryagin's principle to compute the convex hull of the set of reachable points at time t. By letting t goes to infinity, this allows one to compute the convex hull of the asymptotic set A_F (defined in Equation (6)), which is a super-set of the Birkhoff centre B_F.

For the SIR model, we use a more direct approach, that is faster and more accurate for a two dimensional system. By integrating the ODE $\dot{x} = F(x, \vartheta_{\max})$, we first compute a point x_0 that is the fixed point of the uncertain model with fixed parameter $\vartheta_{\max}$. We then compute a trajectory $\mathbf{x}_1$ by integrating the ODE $\dot{x} = F(x, \vartheta_{\min})$ starting in x_0 and a trajectory $\mathbf{x}_2$ by integrating the ODE $\dot{x} = F(x, \vartheta_{\max})$ starting in $\mathbf{x}_1(\infty)$. The two curves delimit a convex region that is included in the Birkhoff centre. We then start from any points on the surface of this region and we look for a value ϑ such that the drift is directed outside this region. If such a value exists, we then enlarge our region by computing a new trajectory with parameter ϑ. We repeat these iterations until no such point on the surface exists. When this region cannot be enlarged, we then obtain the Birkhoff centre of the differential inclusion. This is so because the drift vector never points outwards the region in any point of its boundary, meaning that no trajectory can escape from it.

In Figure 3, we compare the Birkhoff centre of the imprecise fluid model with the one of the uncertain fluid model. The steady-state of the uncertain fluid model is shown in red while the steady-state of the imprecise model corresponds to the whole convex region surrounded by the blue curve. We observe that the steady-state of the uncertain model is strictly included in the one of the imprecise model. Moreover, there are some points of the Birkhoff centre of the imprecise model for which X_S is smaller and X_I is larger that any stationary point of the uncertain model.

D. Comparison with the differential-hull approximation

It should be noted that the algorithms based on the Pontryagin's maximum principle provide an exact numerical method to compute the transient and steady-state behaviour of these systems. On the contrary, the differential-hull approximation, introduced in Section IV-B provides a reasonably accurate results when the range of the parameter θ is small but its accuracy is very poor when the range of parameters increases.

These facts are illustrated in Figure 5 and Figure 4. In Figure 4, we plot the solutions as a function of time. We observe that when the range of possible ϑ is small ($\vartheta_{\max} = 2$), then the differential hull approximation is quite accurate. When $\vartheta_{\max}$ grows, the approximation is less and less accurate. For example, for $\vartheta_{\max} = 5$, the approximation is that $X_I(t) \in$

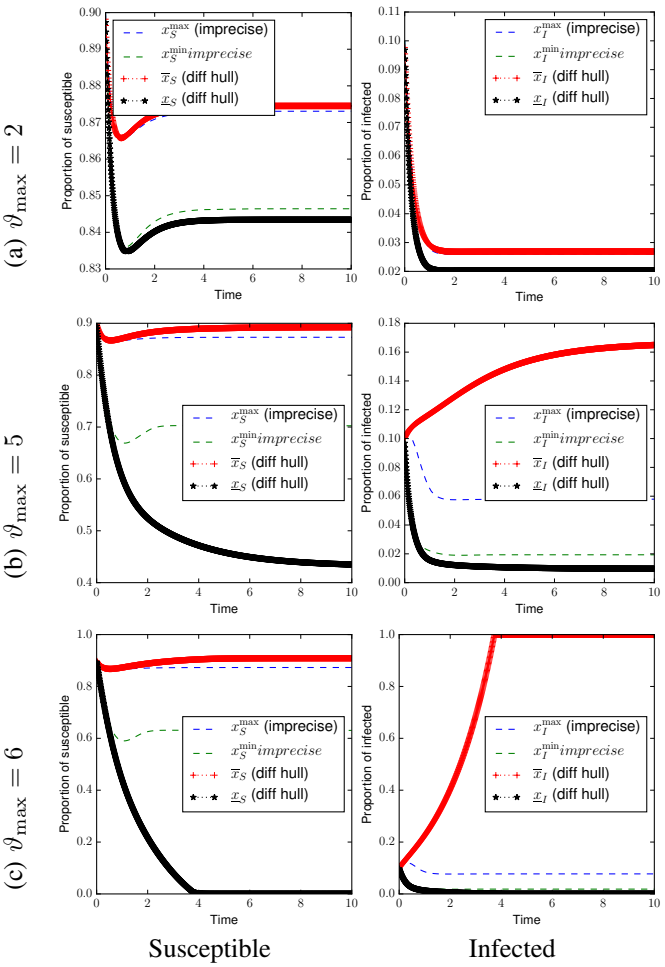


Fig. 4: Evolution of the proportion of susceptible (left) or infected (right) as a function of time for three values of $\vartheta_{\max}$. We compare the differential hull approximation (in dashed line) with the imprecise model (we set $\vartheta_{\min} = 1$).

$[.02, 1.17]$. When $\vartheta_{\max} = 6$, the approximation is trivial for $t \geq 4$, for which we have $\underline{X_I}(t) = 0 \leq X_I(t) \leq \overline{X_I}(t) = 1$.

In Figure 5, we compare the possible steady-states for the imprecise model and the uncertain model with the bounds obtained by the differential-hull approach. The differential-hull approach provides a rectangular approximation for the steady-state distribution. As for the transient regime, the differential-hull approximation deteriorates non linearly in $\vartheta_{\max}$: it is very accurate for $\vartheta_{\max} = 2$ or 3 but very loose for $\vartheta_{\max} = 5$ (and trivial for $\vartheta_{\max} \geq 6$ – not shown on the figure).

E. Comparison between the mean field approximation and stochastic simulations

Theorem 3 guarantees that when N is large, for any control policy θ, the stationary measure of the stochastic system concentrates on a subset of the Birkhoff centre as N goes to infinity. In this section, we illustrate this result by comparing a stochastic simulation of an imprecise Markov population process of the SIR model for a finite N with the limiting regime. We implemented a simulator of the SIR model

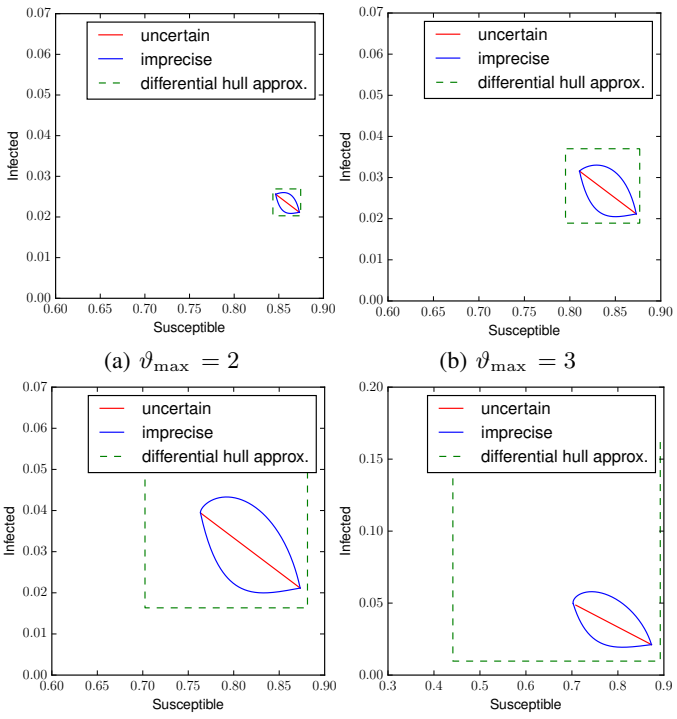


Fig. 5: Comparison of the possible steady-state of the imprecise model, the uncertain and the differential hulls approximation (we set $\vartheta_{\min} = 1$).

with two different control functions $\theta(t)$ that are piece-wise constant:

(a) A function θ^1 that oscillates between $\vartheta_{\min}$ and $\vartheta_{\max}$ according to the following rule: if $\theta^2(t) = \vartheta_{\max}$ and $X_S(t) < .5$, then $\theta^2(t)$ switches to $\vartheta_{\min}$. Conversely, if $\theta^2(t) = \vartheta_{\min}$ and $X_S(t) > 0.85$, then $\theta^2(t)$ switches to $\vartheta_{\max}$.

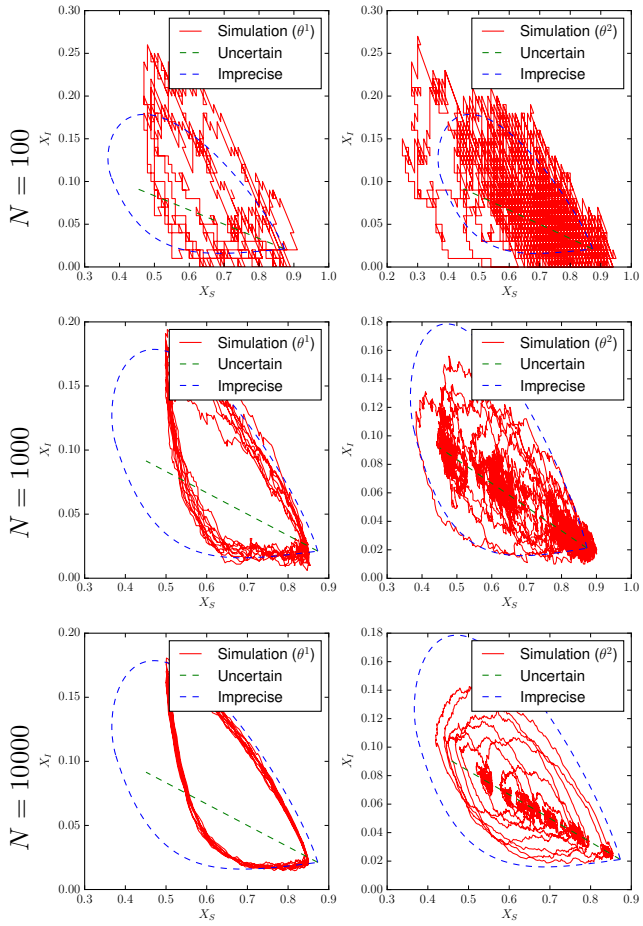
(b) A function θ^2 that jumps to a new value with a rate $5X_I(t)$. The new value is picked uniformly on $[\vartheta_{\min}, \vartheta_{\max}]$.

These functions are chosen as they induce large variations of ϑ, forcing large oscillations of the stochastic system. The control function θ^1 induces an almost-periodic behaviour of the mean field approximation. The function θ^2 is more random.

A sample path in the steady state regime of the stochastic system for $N = 100$, $N = 1000$ and $N = 10000$ is depicted on Figure 6, along with the Birkhoff centre. We observe that for $N \geq 1000$, the stationary behaviour of the stochastic system essentially remains inside of the Birkhoff centre for both policies. The inclusion tends to be exact as N goes to infinity.

VI. Illustrative Example: A Generalised Processor Sharing Model

In this section we consider another example from queueing theory, namely a closed tandem queueing network with two classes of jobs consisting of a queue which serves with GPS



(a) policy θ^1 (b) policy θ^2

Fig. 6: Comparison of simulations of the stochastic system where θ varies with one of the two policies (a) and (b) of Section V-E with the Birkhoff centre of the differential inclusion (in blue). As N grows, the simulation gets included in the Birkhoff centre.

discipline and a delay (i.e., infinite server) station [43], [44]. We consider an imprecise scenario for which the rate of jobs creation belongs to an interval $[\lambda^{\min}, \lambda^{\max}]$. This example serves two purposes. The first is to demonstrate that our framework can be used to perform a robust optimisation of the system parameters. The second is to show in such systems, the precise distribution of the arrival process of jobs is fundamental: When jobs are created according to a Poisson process, the imprecise and uncertain case are similar. This is not true for more general arrival processes for which having a varying arrival rate can lead to a higher queue length than setting a constant $\lambda = \lambda^{\max}$.

A. The model

We consider a system in which N applications are treated by a single machines. There are two types of applications: There are N_1 applications of type 1 and $N_2 = N - N_1$ applications of type 2. Each application sends periodically a job to the machine (the jobs creation process will be described later). The jobs created by the applications of type i are called the jobs of type i. We assume that the size of a job is exponentially distributed with mean $1/\mu_i$ and that the capacity of the machine is C (*i.e.*, on average the machines takes $1/(\mu_i C)$ unit time to serve a job of type i).

Let Q_i be the number of jobs of type $i \in \{1, 2\}$ that are waiting to be completed by the machine. The machine serves the jobs by using a GPS policy. This means it uses a fraction $\phi_i/(\phi_1 Q_1 + \phi_2 Q_2)$ of its time to each job of type i, where $\phi_i > 0$ is a weight. The weight is chosen so as to give priority to a type of jobs. We refer to [44] for a justification and an in-depth analysis of the pertinence of this model in the context of cloud systems, where each application represents one virtual machine or one process running on the same physical machine.

An application that just sent a job waits until the job is completed. Once the job is received, the application becomes idle. It waits a time that we assume is randomly distributed and then sends a new job. In the numerical evaluation, we compare two possible distributions for the time between jobs:

- (Poisson) an application of type i waits an exponentially distributed time of mean $1/\lambda'_i$.
- (MAP) an application of type i waits an exponentially distributed time of mean $1/a_i$ before being active. An active application then sends a job after a time that is exponentially distributed with mean $1/\lambda_i$

We refer to the first scenario as "Poisson arrivals" and to the second one as "MAP" (Markov arrival process).

B. Population model

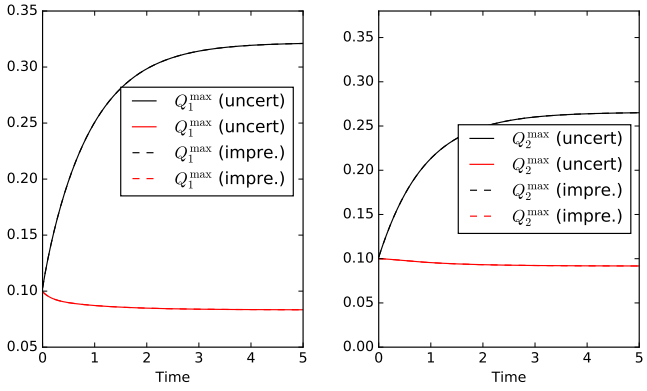
This models fits in our framework. For the Poisson arrival scenario, we describe this system by four variables, Q_i, $i = 1, 2$, the GPS queue length, and D_i, $i = 1, 2$, the number of applications that are waiting before sending a new job. We normalise the variables D_i and Q_i to be in $[0, 1]$, such that $D_i + Q_i = 1$. Transitions of the system are the creation of a job of type i, bringing the system to state $D_i - 1/N_i, Q_i + 1/N_i$, at rate $\lambda_i D_i$, and the service of an job of type i, changing state to $D_i + 1/N_i, Q_i - 1/N_i$. This happens at rate $\mu_i \frac{\phi_i N_i Q_i}{\phi_1 N_1 Q_1 + \phi_2 N_2 Q_2} C$.

The MAP scenario is modelled similarly by adding two other variables E_1 and E_2 to count the applications for which the job was completed but that are not yet active.

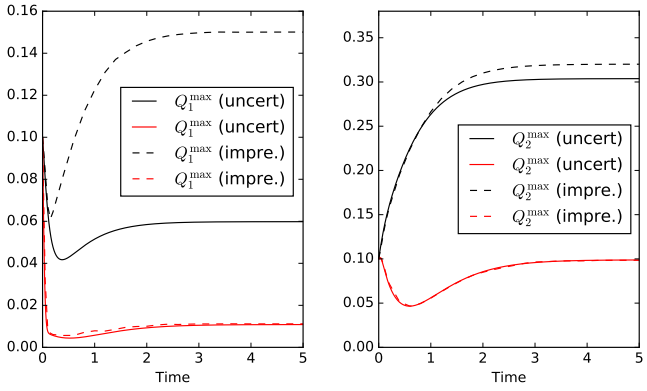
C. Numerical analysis

We consider now a scenario in which the arrival rate is not fixed, but rather an imprecise parameter, varying measurably in $[\lambda_i^{\min}, \lambda_i^{\max}]$. We computed the maximal values of $Q_1(t)$ and $Q_2(t)$ as a function of time for the imprecise and the uncertain model. The results are reported in Figure 7. The parameters of the simulations are $\mu_1 = 5$, $\mu_2 = 1$, $\phi_1 = \phi_2 = 1$, $\lambda_1^{\min} = 1$, $\lambda_1^{\max} = 7$, $\lambda_2^{\min} = 2$, $\lambda_2^{\max} = 3$, $a_1 = 1$ and $a_2 = 2$. We set $\lambda_i'^{\min} = 1/(1/a_i + a/\lambda_i^{\min})$ and $\lambda_i'^{\max} = 1/(1/a_i + a/\lambda_i^{\max})$ which are such that the average time that an application waits in the Poisson and MAP scenario are equal. In each case, the initial state is $Q_1(0) = Q_2(0) = 0.1$.

We make two observations. First, for the Poisson scenario, the uncertain and the imprecise model give the same upper and lower bound on the queue length. This confirms the intuition



(a) Poisson arrivals



(b) Markov arrival process

Fig. 7: GPS model: Maximal queue length as a function of time for the uncertain and the imprecise model. This Figure is to be compared with Figure 1.

that the higher is the arrival rate λ, the more congested is the system. A more surprising fact is that in the MAP case, the maximal queue length is significantly larger than in the imprecise scenario. This fact is counter-intuitive, yet it is a consequence of the delay introduced by the time that the application take to activate. As for the SIR example, these upper and lower bounds were computed by using Pontryagin's maximum principle.

Our imprecise approximation can also be used to robustly tune controllable parameters of our system. We consider a scenario in which we want to set the resource allocation weights ϕ_1 and ϕ_2 in order to minimise the maximum total queue length: $\bar{Q}(t) = \max_\theta(Q_1^\theta(t) + Q_2^\theta(t))$. Using the Pontryagin's principle, we can compute $\bar{Q}(t)$ and minimise it numerically with respect to the weights. As only the ratio of ϕ_1/ϕ_2 plays a role, we fixed the parameter $\phi_2 = 1$. In turns out that the total queue length is a convex function of ϕ_1 and that the minimum value of $\bar{Q}(t)$ is obtained for $\phi_1 = 9.0\phi_2$.

VII. Conclusions

We presented imprecise population processes, which naturally capture the unavoidable uncertainty and imprecision inherent in any (stochastic) model of complex phenomena. In order to analyse efficiently these models in the large population regime, we proved mean field theorems in terms of differential inclusions. We then discussed how to numerically analyse such limit systems, and show the approximation at work in a model of epidemic spreading and of generalised processor sharing.

Future work include a deeper investigation of the Pontraygin method also for steady state computation and in combination with (non-linear) templates, to provide tight bounds to the solutions of a differential inclusion. We will also release an implementation and test the approach on larger models, to properly understand its scalability.

Acknowledgment

This work is supported by the EU project QUANTICOL, 600708.

References

[1] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*. Springer-Verlag, 2000.

[2] Y. C. Stamatiou, P. G. Spirakis, T. Komninos, and G. Vavitsas, *Computer Network Epidemics: MODELS and Techniques for Invasion and Defense*. CRC Press, Inc., 2012.

[3] L. Mari, E. Bertuzzo, L. Righetto, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo, "Modelling cholera epidemics: The role of waterways, human mobility and sanitation," *Journal of The Royal Society Interface*, vol. 9, no. 67, pp. 376–388, 2011.

[4] L. Bortolussi, D. Milios, and G. Sanguinetti, "Smoothed Model Checking for Uncertain Continuous Time Markov Chains," *Information and Computation*, 2015.

[5] M. Benaim and J.-Y. Le Boudec, "A class of mean field interaction models for computer and communication systems," *Performance Evaluation*, vol. 65, no. 11, pp. 823–838, 2008.

[6] J. Aubin and A. Cellina, *Differential Inclusions*. Springer-Verlag, 1984.

[7] E. Todorov, "Optimal control theory," *Bayesian brain: Probabilistic approaches to neural coding*, pp. 269–298, 2006.

[8] J.-P. Aubin and G. D. Prato, "The viability theorem for stochastic differential inclusions 2," *Stochastic Analysis and Applications*, vol. 16, no. 1, pp. 1–15, 1998.

[9] M. Kisielewicz, *Stochastic differential inclusions and applications*. Springer, 2013.

[10] D. Škulj, "Discrete time markov chains with interval probabilities," *International journal of approximate reasoning*, vol. 50, no. 8, pp. 1314–1329, 2009.

[11] T. Kurtz, "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes," *Journal of Applied Probability*, vol. 7, pp. 49–58, 1970.

[12] T. Kurtz and S. Ethier, *Markov Processes - Characterisation and Convergence*. Wiley, 1986.

[13] M. Tschaikowski and M. Tribastone, "Approximate reduction of heterogenous nonlinear models with differential hulls," *Automatic Control, IEEE Transactions on*, 2015.

[14] A. Kolesnichenko, P.-T. de Boer, A. Remke, and B. R. Haverkort, "A logic for model-checking mean-field models," in *Dependable Systems and Networks (DSN), 2013 43rd Annual IEEE/IFIP International Conference on*, IEEE, 2013, pp. 1–12.

[15] L. Bortolussi, J. Hillston, D. Latella, and M. Massink, "Continuous approximation of collective systems behaviour: A tutorial," *Performance Evaluation*, vol. 70, no. 5, pp. 317–349, 2013.

[16] L. Bortolussi, "Hybrid Limits of Continuous Time Markov Chains," in *Proceedings of the 2011 Eighth International Conference on Quantitative Evaluation of Systems (QEST)*, 2011, pp. 3–12.

[17] N. Gast and B. Gaujal, "Markov chains with discontinuous drifts have differential inclusion limits," *Performance Evaluation*, vol. 69, no. 12, pp. 623–642, 2012.

[18] G. Roth and W. H. Sandholm, "Stochastic approximations with constant step size and differential inclusions," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 525–555, 2013.

[19] C. Baier, H. Hermanns, J.-P. Katoen, and B. R. Haverkort, "Efficient computation of time-bounded reachability probabilities in uniform continuous-time Markov decision processes," *Theor. Comput. Sci.*, vol. 345, no. 1, pp. 2–26, 2005.

[20] M. Beccuti, E. Amparore, S. Donatelli, D. Scheftelowitsch, P. Buchholz, and G. Franceschinis, "Markov decision petri nets with uncertainty," in *Proceedings of 12th European Workshop Computer Performance Engineering, EPEW*, ser. Lecture Notes in Computer Science, vol. 9272, 2015, pp. 177–192.

[21] R. Givan, S. Leach, and T. Dean, "Bounded-parameter markov decision processes," *Artificial Intelligence*, vol. 122, no. 1-2, pp. 71 –109, 2000.

[22] C. Fricker and N. Gast, "Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity," *EURO Journal on Transportation and Logistics*, pp. 1–31, 2014.

[23] M. Benaïm, J. Hofbauer, and S. Sorin, "Stochastic approximations and differential inclusions," *SIAM Journal on Control and Optimization*, vol. 44, no. 1, pp. 328–348, 2005.

[24] O. Maler, "Computing reachable sets: An introduction," Tech. rep. French National Center of Scientific Research. www-verimag. imag. fr/maler/Papers/reach-intro. pdf, Tech. Rep., 2008.

[25] B. De Schutter, W. Heemels, J. Lunze, and C. Prieur, "Survey of modeling, analysis, and control of hybrid systems," in *Handbook of Hybrid Systems Control–Theory, Tools, Applications*, 2009, pp. 31–55.

[26] A. B. Kurzhanski and P. Varaiya, "On ellipsoidal techniques for reachability analysis.," *Optimization methods and software*, vol. 17, no. 2, 2002.

[27] A. Girard, C. Le Guernic, and O. Maler, "Efficient computation of reachable sets of linear time-invariant systems with inputs," in *Proceedings of HSCC 2006*, ser. LNCS, 2006.

[28] G. Alefeld and G. Mayer, "Interval analysis: Theory and applications," *Journal of computational and applied mathematics*, vol. 121, no. 1, pp. 421–464, 2000.

[29] X. Chen, E. brahm, and S. Sankaranarayanan, "Flow*: An analyzer for non-linear hybrid systems," in *Computer Aided Verification*, Springer, 2013, pp. 258–263.

[30] N. Ramdani and N. S. Nedialkov, "Computing reachable sets for uncertain nonlinear hybrid systems using interval constraint-propagation techniques," *Nonlinear Analysis: Hybrid Systems*, vol. 5, no. 2, pp. 149–162, 2011.

[31] S. Gao, S. Kong, and E. M. Clarke, "Dreal: An SMT solver for nonlinear theories over the reals," in *Automated Deduction?CADE-24*, Springer, 2013, pp. 208–214.

[32] L. De Moura and N. Bjrner, "Z3: An Efficient SMT Solver," en, in *Tools and Algorithms for the Construction and Analysis of Systems*, ser. Lecture Notes in Computer Science 4963, C. R. Ramakrishnan and J. Rehof, Eds., Springer Berlin Heidelberg, 2008, pp. 337–340.

[33] O. Botchkarev and S. Tripakis, "Verification of hybrid systems with linear differential inclusions using ellipsoidal approximations," in *Hybrid Systems: Computation and Control*, Springer, 2000, pp. 73–88.

[34] G. e. a. Frehse, "Spaceex: Scalable verification of hybrid systems," in *Proceedings of CAV 2011*, ser. LNCS, vol. 6806, 2011.

[35] T. Dang, C. Le Guernic, and O. Maler, "Computing reachable states for nonlinear biological models," *Theor. Comput. Sci.*, vol. 412, no. 21, 2011.

[36] P. Zgliczynski and T. Kapela, "Lohner algorithm for perturbation of odes and differential inclusions," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 11, no. 2, pp. 365–385, 2009.

[37] A. Donzé and O. Maler, "Systematic simulation using sensitivity analysis," in *Proceedings of HSCC 2007*, 2007.

[38] L. Bortolussi and G. Sanguinetti, "A Statistical Approach for Computing Reachability of Non-linear and Stochastic Dynamical Systems," in *Quantitative Evaluation of Systems*, ser. LNCS 8657, G. Norman and W. Sanders, Eds., springer, 2014, pp. 41–56.

[39] A. Bhatia and E. Frazzoli, "Incremental search methods for reachability analysis of continuous and hybrid systems," *Hybrid Systems: Computation and Control*, pp. 451–471, 2004.

[40] T. Dang and T. Dreossi, "Falsifying oscillation properties of parametric biological models," in *Proceedings of HSB 2013*, ser. EPTCS, vol. 125, 2013.

[41] T. Dang and T. Nahhal, "Coverage-guided test generation for continuous and hybrid systems," *Formal Methods in System Design*, vol. 34, no. 2, 2009.

[42] N. Ramdani, N. Meslem, and Y. Candau, "A hybrid bounding method for computing an over-approximation for the reachable set of uncertain nonlinear systems," *IEEE Trans. Automat. Contr.*, 23522364, 10 2009.

[43] G. Iacobelli and M. Tribastone, "Lumpability of fluid models with heterogeneous agent types," in *Proceedings of the 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, IEEE, 2013, pp. 1–11.

[44] J. Anselmi and I. Verloop, "Energy-aware capacity scaling in virtualized environments with performance guarantees," *Performance Evaluation*, vol. 68, no. 11, pp. 1207–1221, 2011.