QUANTICOL

A Quantitative Approach to Management and Design of Collective and Adaptive Behaviours

http://www.quanticol.eu

TR-QC-07-2015

Mean Field Approximation of Imprecise Population Processes

Revision: 1.0; September 25, 2015

Author(s): Luca Bortolussi (CNR); Nicolas Gast (INRIA)

Publication date: Sept 25, 2015

Funding Scheme: Small or medium scale focused research project (STREP)
Topic: ICT-2011 9.10: FET-Proactive 'Fundamentals of Collective Adaptive Systems' (FOCAS)
Project number: 600708

Coordinator: Jane Hillston (UEDIN)
e-mail: Jane.Hillston@ed.ac.uk
Fax: +44 131 651 1426

Part. no.	Participant organisation name	Acronym	Country
1 (Coord.)	University of Edinburgh	UEDIN	UK
2	Consiglio Nazionale delle Ricerche – Istituto di Scienza e Tecnologie della Informazione "A. Faedo"	CNR	Italy
3	Ludwig-Maximilians-Universität München	LMU	Germany
4	Ecole Polytechnique Fédérale de Lausanne	EPFL	Switzerland
5	IMT Lucca	IMT	Italy
6	University of Southampton	SOTON	UK
7	Institut National de Recherche en Informatique et en Automatique	INRIA	France

COOPERATION

Abstract

We consider stochastic population processes in presence of uncertainty, originating from lack of knowledge of parameters or by unpredictable effects of the environment. We set up a formal framework for imprecise population 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 for an infinite population size, proving it is given by a differential inclusion constructed from the (imprecise) drift. We discuss also the steady state behaviour of such a mean field approximation. 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. In the paper, we discuss separately the simpler case of models with a more constrained form of imprecision, in which lack of knowledge on parameter values allows us to assess that they belong to a given interval, albeit being constant in time. Such uncertain population models are amenable of simpler forms of analysis. The theoretical results are accompanied by an in-depth analysis of a simple epidemic model.

Contents

1 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 at most can be estimated from data 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 [2] or of a malware spreading in a computer network [35]. 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 epidemics [2]. However, statistical estimation can never provide an exact value, as we have at disposal only a finite, and often incomplete, amount of data. Moreover, the infection rate itself may depend on environmental factors, which can change arbitrarily in time. For instance, in cholera spreading [30], 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 bike sharing systems. The arrival rate of clients in a bike station cannot be assumed fixed during the day, but it will depend on the hour. The precise form of such a variation 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.

Hence, uncertainty reflects in at least two ways in models of complex systems:

1. Some parameters ϑ can depend on features of the environment external to the model. We may not know the precise value of ϑ, because we cannot measure these features. Furthermore, if ϑ depends on quantities like temperature, PH, atmospheric weather, light intensity, and so on, it may be subject to variations during the time horizon T of interest, so that considering it as a fixed parameter may be 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 just 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. that $\vartheta_t \in \Theta$. We call this the *imprecise* scenario.

2. In a simpler scenario, a parameter ϑ is assumed fixed, but its precise value not known precisely. In this case, we just 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.

We consider models of large populations of interacting agents, formalised in terms of Continuous Time Markov Chains. Usually, these models are fully specified, meaning that all parameters have a fixed and constant value. This is however striding with the intrinsic uncertainty of the world. In this paper, we will overcome this limitation by introducing a novel class of population CTMC models, which we call *Imprecise Population Processes*. 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 evolution 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. In this case, the class of population models so obtained is considerably simpler, resulting in the so called Uncertain Continuous Time Markov Chains, see [11].

When populations are large, as is often the case in epidemiology, biology, telecommunications, a direct analysis of the stochastic model is unfeasible, even by simulations and statistics. A viable alternative is to construct the so-called mean field approximation [10], which provides a deterministic description of the expected behaviour to which stochastic trajectories converge in the limit of infinite populations. The complexity of the analysis is 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, [3, 23]), both for transient and steady-state behaviour (when this is meaningful) and investigate the computational gains obtained in this way. With respect to classic mean field, the presence of imprecision explodes 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 method to bound the solutions of DI using control theoretic tools, namely the Pontraygin principle [36, Section 3].

The paper is organised as follows: Section 2 discusses Imprecise Markov Chains and the evolution of their probability mass in terms of differential inclusions. Section 3 introduces imprecise population models and proves their mean field behaviour. Section 4 discusses computational methods for the analysis of the differential inclusion mean field limits. Section 5 works out in detail an example about epidemic spreading. The final discussion is in Section 6.

Related work

Following the terminology of [4], 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 similarly to the notion on stochastic differential inclusion studied in [4, 25], 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 [34], that considers Markov chain with interval probabilities. The key contribution of our paper with respect to [34] 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 [27, 26]. Mean field based analysis, for the transient and the steady state, have been applied in performance modelling, systems biology, epidemiology [7]. For a gentle introduction, see [10]. Classic mean field results require (Lipschitz) continuity of rate functions, but more general theorems can be proved for piecewise smooth rates [9, 23] or even general discontinuous functions by using differential inclusions [23].

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 [23] and can be seen as corollaries of [33].

2 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 model, and the subclass of uncertain Continuous Time Markov Chains. 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.

2.1 Imprecise and uncertain Markov chains

We consider a stochastic process $\mathbf{X} = (X_t)_{t\geq 0}$, adapted to a filtration[1] $\mathcal{F}$, that takes value in a state space $\mathbf{E} \subset \mathbb{R}^d$. 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, and we consider a set of transitions kernel on $\mathbf{E}$, parametrized by $\vartheta \in \Theta$: for each $\vartheta \in \Theta$, Q^ϑ is a transition kernel, *i.e.* 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 continous time Markov chain is a stochastic process* $\mathbf{X}$ *together with a* $\mathcal{F}_t$*-adapted process* θ *such that for all* $t \geq 0$*:*

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

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* $t \geq 0$*:*

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

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 process 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) [5], the main difference being that in MDPs, the emphasize is on finding a policy that maximizes 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 characterize the set of possible behaviours.

Example. For illustrative purposes, we consider a simple example of a bike sharing system. We describe the number of bikes present in a single station, so that the state of the CTMC is given by $X_t \in \{0, \ldots, N\}$, 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 $[\theta_a^{min}, \theta_a^{max}]$. 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 $[\theta_r^{min}, \theta_r^{max}]$, 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 $\theta_a(t)$ and $\theta_r(t)$ 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,

[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.

we could assume these values are unknown but constant in time, for example if we are considering the dynamics of the station in a restricted time frame. The return rate $\theta_r(t)$ can also be taken as a function of the number of free slots in the station, $N - X_t$, taken as an indicator of the amount of bikes currently circulating in the system.

2.2 Differential inclusions

Differential inclusions (DI) are a generalization of differential equations, and provide the natural mathematical tool to describe the transient evolution of probability mass of an imprecise CTMC. Furthermore, as we will see in the next section, under a proper scaling, the behavior of 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 [3] 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 vector $F(x) \subset \mathbb{R}^d$. 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 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, we denote by $S_{F,x}$ (or S_x if there is no ambiguity) the set of solutions of $\dot{x} \in F(x)$ that start in x. Note that the set $S_{F,x}$ 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 behavior as t goes to infinity. As the time grows, a solution $\mathbf{x}$ of a differential inclusion can have a chaotic behavior and can oscillate around many points. For a given starting point x, we define the limit set $L_F(x)$ as the closure of the set of points that are accumulation points of at least one solution of the differential inclusion starting in x, *i.e.*,

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

As in [33], we call the Birkhoff center 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 center 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 center 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 center 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 [7]).

2.3 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_t, 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_t, 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 unpractically large in most interesting scenarios, especially for population models, see Section 3.

2.4 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. *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. We call this sum the imprecise drift of the imprecise CTMC. It is a function $f : E \times \Theta \to \mathbb{R}^d$, defined by*

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

3 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 Birkoff center.

3.1 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 $(X^N)_N$ 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 $Q^{N,\theta}$.

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:* $\sup_{x \in E^N, \vartheta \in \Theta} Q^{N,\vartheta}_{xx} < \infty$

(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 [10]. The idea is to specify the dynamics by a set of possible transitions or events, providing their rate, as a function of 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 5 for examples.

Example. A simple example of a population model is provided by a simple 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 $N\theta_a(t)$, 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 $N\theta_r(t)$, 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.

3.2 Mean field limit (finite time-horizon)

Let $f^N(x, \vartheta) = \sum_{y \in E^N} Q_{xy}^{N,\vartheta}(y - x)$ 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:

$$\begin{aligned} F(x) &= \overline{\lim_{N\to\infty}} \bigcup_{\vartheta\in\Theta} \left\{ f^N(x, \vartheta) \right\}, \\ &= \left\{ y \in \mathbb{R}^d \text{ such that } \exists \phi(N), \lim_{N\to\infty} \phi(N) = \infty \wedge \lim_{N\to\infty} f^{\phi(N)}(x^{\phi(N)}, \vartheta^{\phi(N)}) = y \right\} \end{aligned} \quad (4)$$

Theorem 1. *Let (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 the differential inclusion $\dot{x} \in F(x)$ starting in x.*

Proof. It can be shown that there exists a deacreasing function $I(N)$ such that the discrete time process $\mathbf{X}^N(kI(N))$ satisfies the definition of a GASP of [33]. Then, the proof follows from [33, Theorem 3.1]. □

Remark. The construction of the imprecise drift generalises the definition given in [23], where the focus was on discontinuous, but precise, rates. The proof of the theorem follows the ides of [6, 23]. In fact, when Θ is restricted to a single value, we obtain again the result of [23].

In the case of an uncertain Markov chain, we define the drift by

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

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.*

An illustration of these results at work is postponed to Section 5.

3.3 Mean field limit (stationary regime)

An uncertain Markov chain does not necessarily have a stationary behavior. 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 a uncertain Markov population process by using the asymptotic reachable set of the differential inclusion, A_F, defined by:

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

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

Theorem 2. *Let* $\mathbf{X}$ *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 characterize the stationary behavior of the stationary measures as N grows. It shows that the sequence of stationary measures μ^N concentrates on the Birkhoff center of the differential inclusion. The Birkhoff center, defined in Section 2.2, 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 center of* F*, defined in Equation* (1)*:* $\mu(B_F) = 1$*.*

Proof. As $\mathbf{X}$ is a GASP, the proof follows from [33, 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 center of* F_ϑ*, defined in Equation* (1)*.*

Remark. Theorem 3 only states that the support of μ is included in the Birkhoff center 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 [33, Definition 3.3]. However, this notion is very 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 center, which, even if less accurate, provides a simpler characterization that can be used numerically.

4 Numerical Methods and Algorithmic Issues

The results of 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 originating population process. The analysis of this class of mean field models, however, is in general considerably more challenging than Ordinary Differential Equations. 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.

4.1 Related Work

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 [29].

The first class of methods manipulates directly sets of states, finitely represented, for instance, as polytopes [19], ellipsoids [28], or zonotopes [24], or by relying on interval arithmetic [1] combined with constraint solving [14, 32], exploiting satisfaction modulo solvers over reals [22, 18]. 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 linear systems, due to their wide diffusion in engineering applications. For this class of systems, efficient methods exist [24, 13]. Tools like SpaceEx [21] 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, as the problem is much more difficult. Here we recall hybridization [16], 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 [14] and [32]. Other methods for integrating differential inclusions rely on error bounding [38].

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 [20], which uses the sensitivity of the system with respect to initial conditions to (approximatively) compute how a set of initial conditions propagates in T units of time under the action of the dynamics. The approach of [12], instead, use statistical methods borrowed from machine learning to perform an inference of the reachable set with statistical error guarantees. Other methods work for more general imprecise limit models; for example, the procedure of [8, 15, 17] constructs an under-approximation of the reachable set using a Monte-Carlo sampling method.

4.2 Differential hull of a differential inclusion

Our first method to compute the set of reachable points by the solutions of a differential inclusion is to construct rectangular bounds for the differential inclusion, *i.e.*, two functions $\underline{x}$ and $\overline{x}$ such that $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ for any solution of the differential inclusion x. In this section, we show that we can construct to differential equations such that $\underline{x}$ and $\overline{x}$ are solutions of a simple differential equations 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 5.4, the bounds provided by this approximation become too large when the possible values of ϑ increase.

Following the definition of [37], we say that a locally Lipschitz-continous 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:

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

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

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

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:*

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

Proof. The result is a direct consequence of Theorem 1 of [31]. □

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

In general, the bounds $\underline{x}$ and $\overline{x}$ are not tight. In this section, we show how to use Pontryagin's maximum principle to compute 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.

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

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

Pontryagin's maximum principle is a set of necessary conditions that the trajectory that attains the maximum should satisfy. Following the description of [36, Section 3] these conditions are the following. If x is a trajectory that maximizes $x_i(T)$, then there exists a *costate trajectory* $\mathbf{p}$ such that $p_i(T) = -1$, $p_j(T) = 0$ for $j \neq i$ and:

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

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

This solution leads to iterative numerical methods that start from an initial costate p, then update p by computing a trajectory x forward in time the x by using (7) and (8) and then computing a new p backward in time by solving the ODE (9).

Remark. The rectangle delimited by $x^{\min}(T)$ and $x_i^{\max}(T)$ provide 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. The algorithm given by the iterations (7-9) can be easily extended to refine the rectangle into any convex template polyhedron by considering the minimization problems $\min_i \sum_i \alpha_i x_i(t)$ for any tuple of coefficients α_i.

5 Illustrative Example: 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 compare the accuracy of the numerical solutions. This example is close to the one used in [37] for heterogeneous systems.

5.1 The SIR model

We consider a system composed of N nodes. Each agent can have three states: susceptible, infected or recovered. We denote by $X_S(t) \in [0,1]$ the proportion of susceptible nodes (and by $X_I(t)$ and $X_R(t)$ the density of infected and recovered nodes). 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. Hence, a susceptible encounters an infected node and becomes infected at rate $\vartheta X_I(t)$.

The transitions of the system are the following: a state X_S, X_I, X_R becomes: $X_S - \frac{1}{N}, I + \frac{1}{N}, X_R$ at rate $N(aX_S + \vartheta X_S X_I)$, it becomes $X_S, X_I - \frac{1}{N}, X_R + \frac{1}{N}$ at rate NbX_I and becomes $X_S + \frac{1}{N}, X_I, X_R - \frac{1}{N}$ at rate NcX_R.

The drift of the system is 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 behavior 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 behavior 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 what follows, we set the parameters equal to $a = 0.1, b = 5, c = 1, \vartheta_{\min} = 1, \vartheta_{\max} = 10$ and the initial conditions are $X_S(0) = 0.7, X_I(0) = 0.3$.

5.2 Reachable sets in finite time

For a fixed parameter ϑ, the drift $f(X_S, X_I, \vartheta)$ is Lipschitz-continous 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 $X_S^\vartheta, X_I^\vartheta$. Although we believe that this ODE does not have a close-form solution, 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 models is included in the set of reachable points by 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.

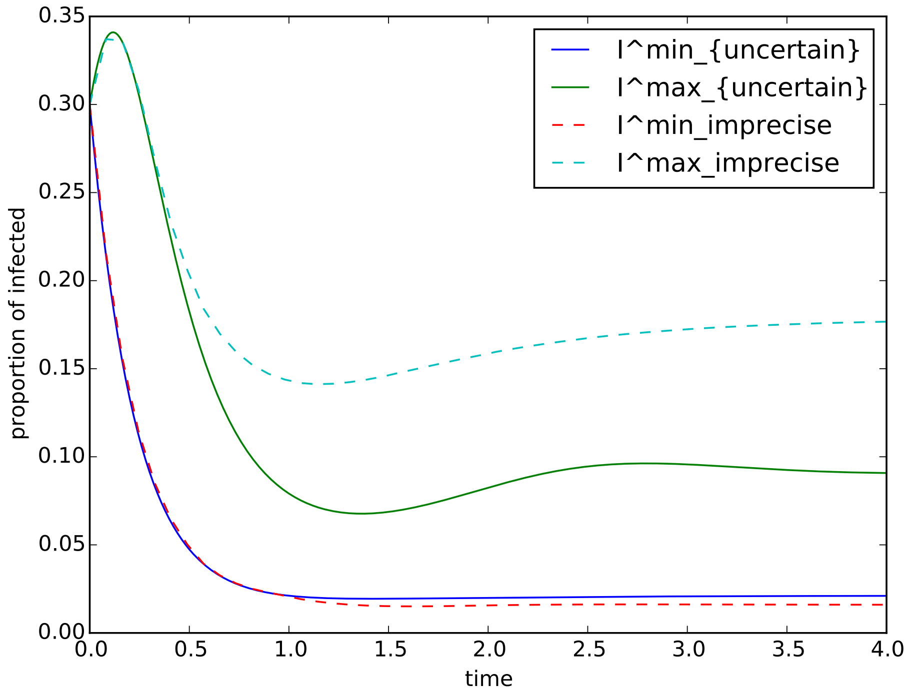


Figure 1: Upper and lower bounds on the number of infected nodes for an imprecise model (dashed lines) and an uncertain model (solid lines)

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

$$\overline{x}_I^{uncertain}(t) = \max_{\vartheta} I_\vartheta(t) \qquad \text{and} \qquad \overline{x}_I^{imprecise} = \max_{(X_S,X_I)\in S_F} X_I(t).$$

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. For the computation $\overline{x}_I^{uncertain}(t)$, we formulate the problem has the optimal control problem which is to find a function $\theta(t)$ that maximizes $X_I(t)$, 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 4.3). 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.

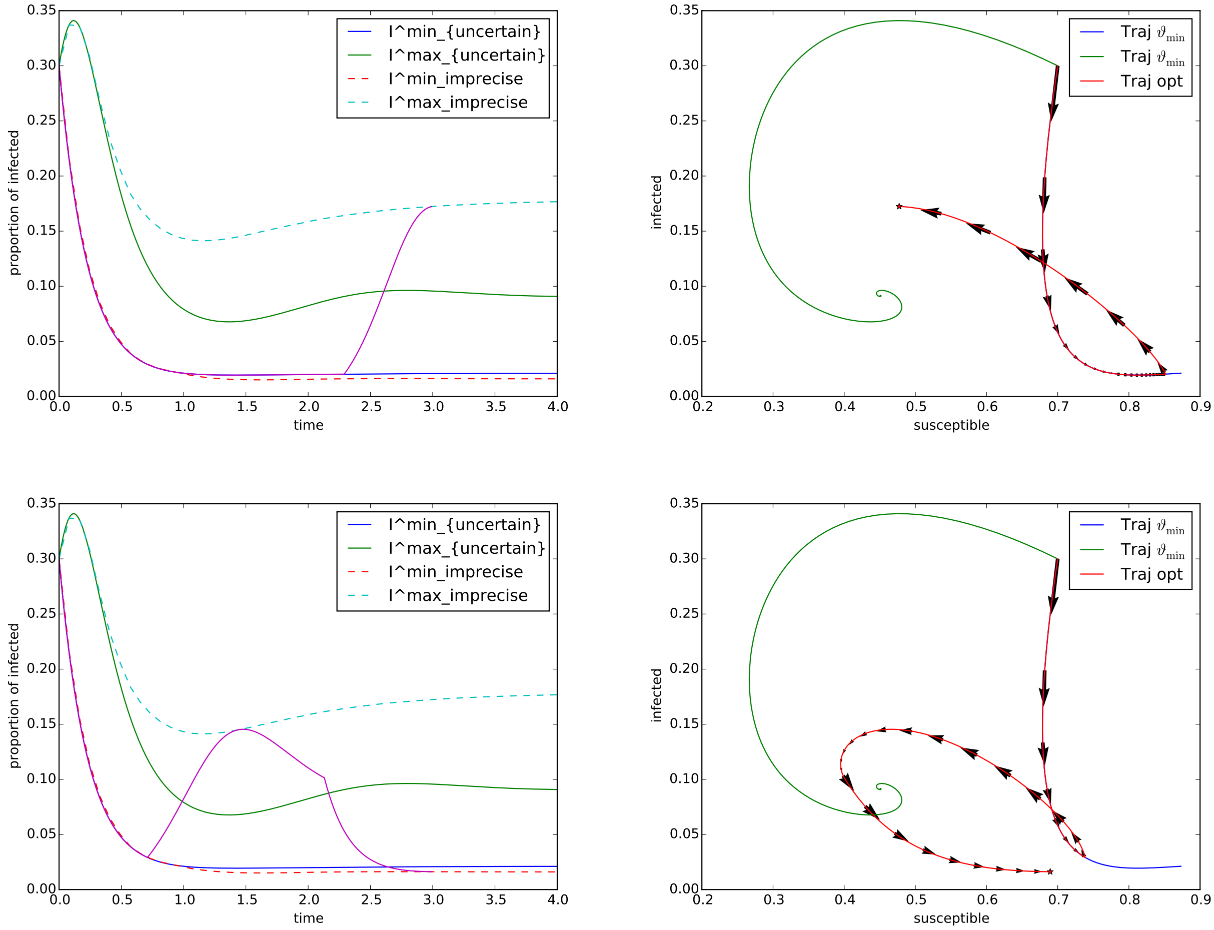


Figure 2: Trajectory that reaches the maximum (top) or minimum (bottom) number of infected node at time $T = 3$. On the left, we plot the number of infected as a function of time. On the right, we plot the number of infected as a function of the number of susceptible.

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 $T = 3$.

For the maximum, the optimal $\theta(t)$ is to set $\theta(t) = \vartheta_{\text{min}}$ for $t < 2.25$ and $\theta(t) = \vartheta_{\text{max}}$ for $t > 2.25$. For the minimum, the optimal $\theta(t)$ is start with $\theta(t) = \vartheta_{\text{min}}$ for $t < 0.7$, then use $\theta(t) = \vartheta_{\text{max}}$ until 2.2 and then use again $\theta(t) = \vartheta_{\text{min}}$.

Also, it should be noted that by using Pontryagin's ideas, we can also compute trajectories that maximizes any utility functions.

5.3 Steady-state regime

Computing the Birkhoff center 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 center 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 $\dot{x} = F(x, \vartheta_{\text{max}})$, we first compute a point x_0 that is the fixed point of the uncertain model with fixed parameter ϑ_{max}. We then compute a trajectory $\mathbf{x}_1$ by integrating the ODE $\dot{x} = F(x, \vartheta_{\text{min}})$ starting in x_0 and a trajectory $\mathbf{x}_2$ by integrating the ODE $\dot{x} = F(x, \vartheta_{\text{max}})$ starting in $\mathbf{x}_1(\infty)$. The two curves delimit a convex region that is included in the

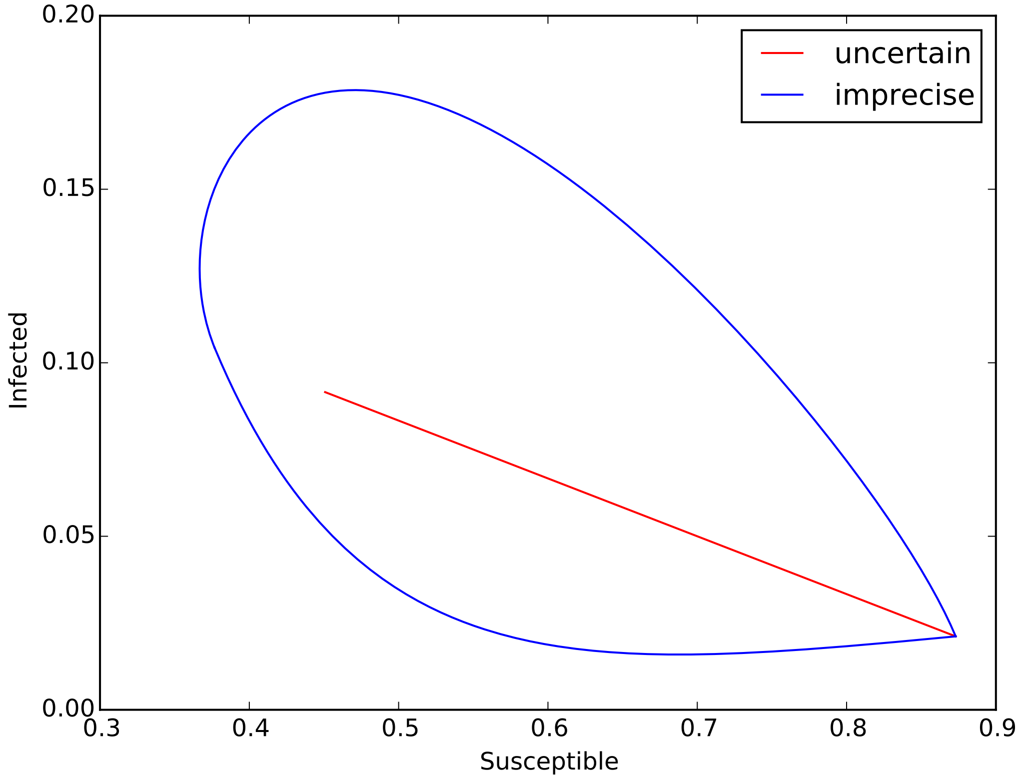


Figure 3: Steady-state regime for the imprecise and uncertain SIR models. The steady-state of the imprecise model is the convex set delimited by the blue region. The steady-state of the uncertain model is on the red line. We set $\vartheta_{\max} = 10\vartheta_{\min}$.

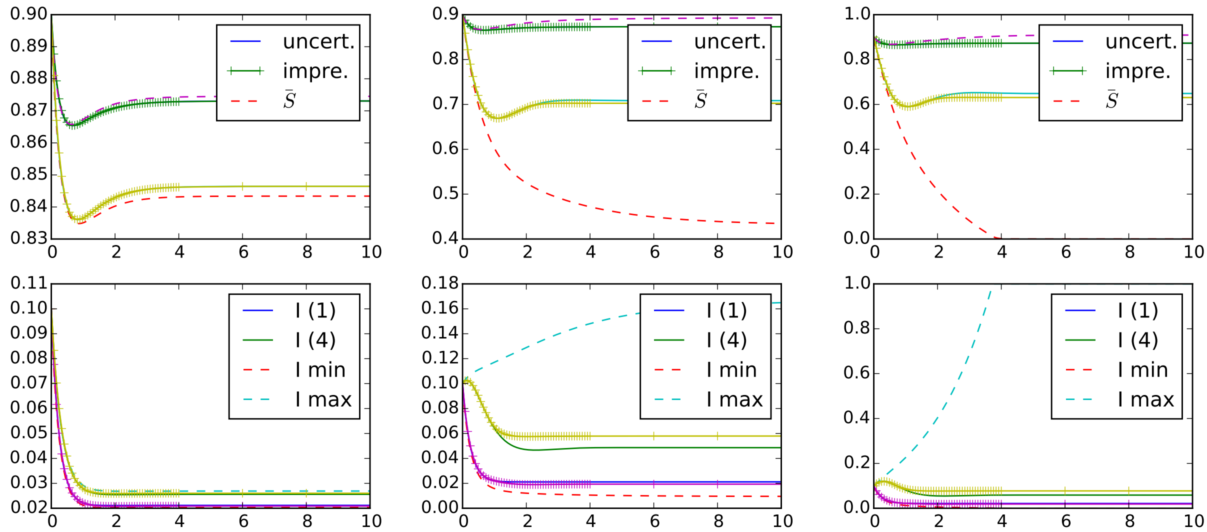
Birkhoff center. We then start from any points on the surface of this region and we look for a value ϑ such that the drift is directed outside this region. If such a value exists, we then enlarge our region by computing a new trajectory. We repeat these iterations until no such point exists. When this region cannot be enlarged, we then obtain a tight over approximation of the Birkhoff center of the differential inclusion. This is so because in no points of the obtained boundary, the drift vector points outwards the region, meaning that no trajectory can escape from it.

In Figure 3, we compare the Birkhoff center 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 center of the uncertain model for which X_S is smaller and X_I is larger that any stationary point of the imprecise model.

5.4 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 behavior of these systems. On the contrary, the differential-hull approximation, introduced in Section 4.2 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 ($\theta_{\max} = 2$), then the differential hull



(a) $\theta_{\max} = 2$ (b) $\theta_{\max} = 5$ (c) $\theta_{\max} = 6$

Figure 4: Evolution of the proportion of susceptible (top) or infected (bottom) as a function of time for three values of $\theta_{\max}$. We compare the differential hull approximation (in dashed line) with the uncertain and imprecise model.

approximation is quite accurate. When $\theta_{\max}$ grows, the approximation is less and less accurate. For example, for $\theta_{\max} = 5$, the approximation is that $X_I(t) \in [.02, 1.17]$. When $\theta_{\max} = 6$, 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 $\theta_{\max}$: it is very accurate for $\theta_{\max} = 2$ or 3 but very loose for $\theta_{\max} = 5$ (and trivial for $\theta_{\max} \geq 6$ – not shown on the figure).

5.5 Accuracy of mean field approximation

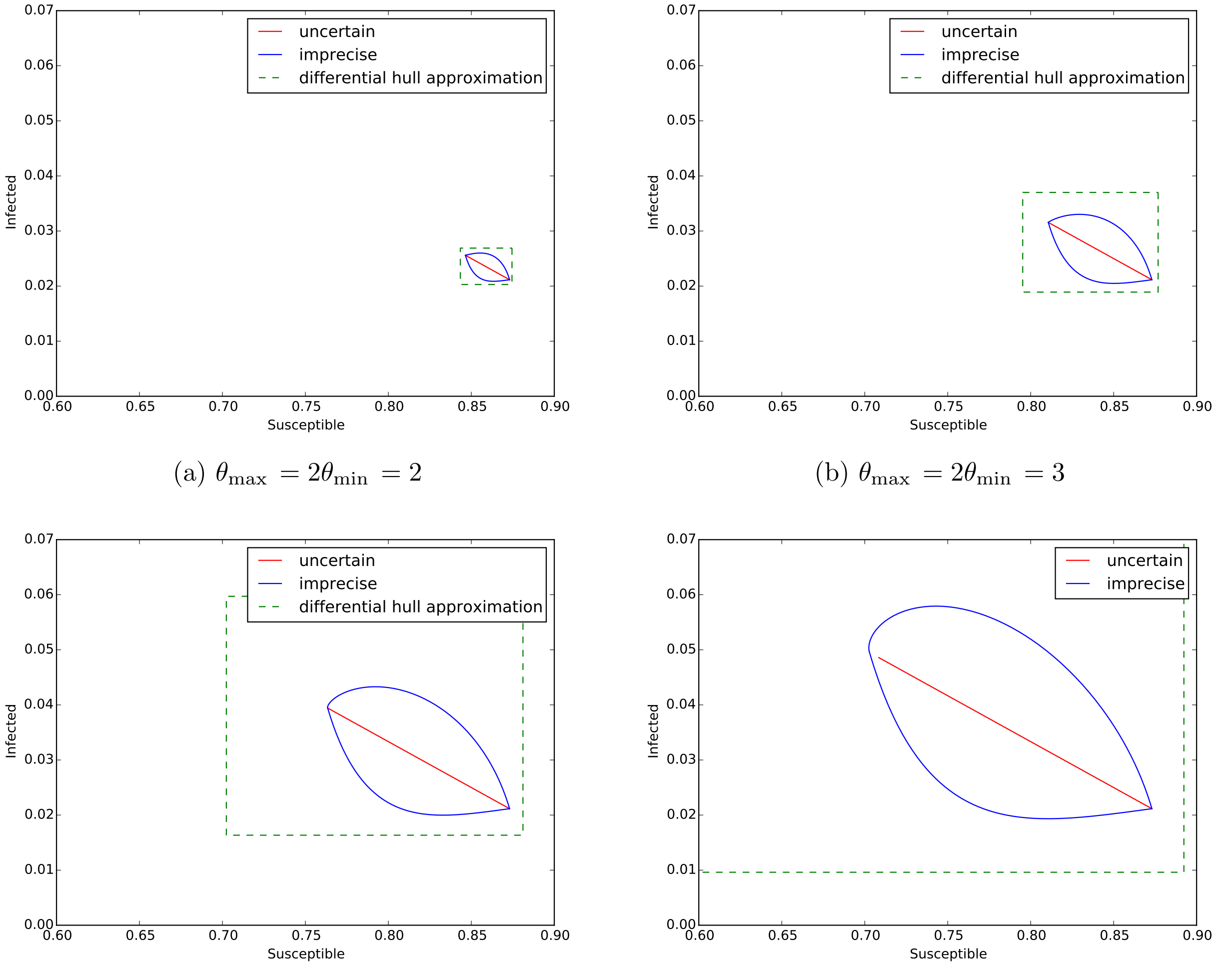
In this section, we compare a stochastic simulation of an imprecise Markov population process of the SIR model for a finite N with the limiting regime. For the simulation, we compare two functions $\theta(t)$ that are piece-wise constant:

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

(b) 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 $\vartheta_{\max}$.

These functions are chosen as they induce a pretty large variation of ϑ, forcing large oscillations in the stochastic regime.

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 center. We observe that for $N \geq 1000$, the stationary behavior of the stochastic system essentially remains inside of the Birkhoff center for both policies. Such inclusion tends to be strict as N goes to infinity.



(a) $\theta_{\max} = 2\theta_{\min} = 2$

(b) $\theta_{\max} = 2\theta_{\min} = 3$

(c) $\theta_{\max} = 2\theta_{\min} = 4$

(d) $\theta_{\max} = 2\theta_{\min} = 5$

Figure 5: Comparison of the possible steady-state of the imprecise model, the uncertain and the differential hulls approximation.

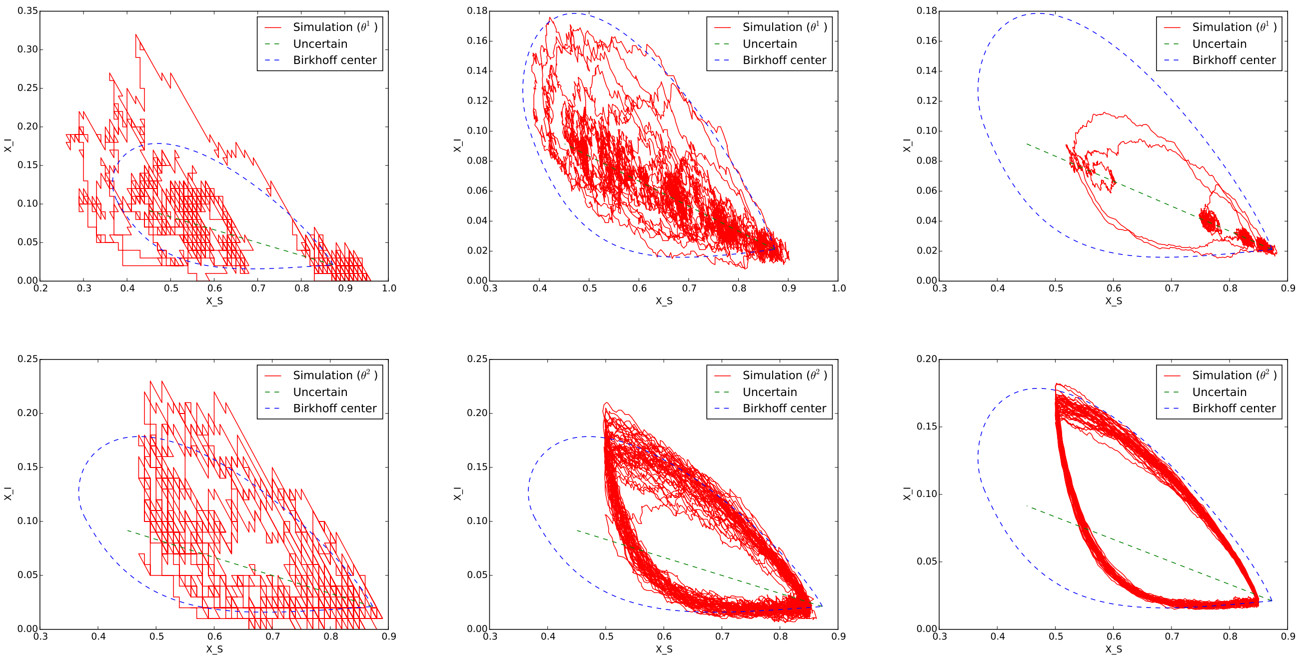
6 Conclusions

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.

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.

References

[1] G. Alefeld and G. Mayer. Interval analysis: theory and applications. *Journal of computational and applied mathematics*, 121(1):421–464, 2000.



$N = 100$ $N = 1000$ $N = 10000$

Figure 6: Comparison of simulations of the stochastic system where θ varies with one of the two policies (a) and (b) of Section 5.5 with the Birkoff center of the differential inclusion (in blue). As N grows, the simulation gets included in the Birkoff center.

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

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

[4] Jean-Pierre Aubin and Giuseppe Da Prato. The viability theorem for stochastic differential inclusions 2. *Stochastic Analysis and Applications*, 16(1):1–15, 1998.

[5] 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.*, 345(1):2–26, 2005.

[6] Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44(1):328–348, 2005.

[7] Michel Benaim and Jean-Yves Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11):823–838, 2008.

[8] Amit Bhatia and Emilio Frazzoli. Incremental search methods for reachability analysis of continuous and hybrid systems. *Hybrid Systems: Computation and Control*, pages 451–471, 2004.

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

[10] L. Bortolussi, J. Hillston, D. Latella, and M. Massink. Continuous approximation of collective systems behaviour: a tutorial. *Performance Evaluation*, 70(5):317–349, 2013.

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

[12] L. Bortolussi and G. Sanguinetti. A Statistical Approach for Computing Reachability of Non-linear and Stochastic Dynamical Systems. In G. Norman and W. Sanders, editors, *Quantitative Evaluation of Systems*, number 8657 in LNCS, pages 41–56. springer, 2014.

[13] Oleg Botchkarev and Stavros Tripakis. Verification of hybrid systems with linear differential inclusions using ellipsoidal approximations. In *Hybrid Systems: Computation and Control*, pages 73–88. Springer, 2000.

[14] Xin Chen, Erika brahm, and Sriram Sankaranarayanan. Flow*: An analyzer for non-linear hybrid systems. In *Computer Aided Verification*, pages 258–263. Springer, 2013.

[15] T. Dang and T. Dreossi. Falsifying oscillation properties of parametric biological models. In *Proceedings of HSB 2013*, volume 125 of *EPTCS*, 2013.

[16] T. Dang, C. Le Guernic, and O. Maler. Computing reachable states for nonlinear biological models. *Theor. Comput. Sci.*, 412(21), 2011.

[17] T. Dang and T. Nahhal. Coverage-guided test generation for continuous and hybrid systems. *Formal Methods in System Design*, 34(2), 2009.

[18] Leonardo De Moura and Nikolaj Bjrner. Z3: An Efficient SMT Solver. In C. R. Ramakrishnan and Jakob Rehof, editors, *Tools and Algorithms for the Construction and Analysis of Systems*, number 4963 in Lecture Notes in Computer Science, pages 337–340. Springer Berlin Heidelberg, 2008.

[19] B. De Schutter, WPMH Heemels, J. Lunze, and Christophe Prieur. Survey of modeling, analysis, and control of hybrid systems. In *Handbook of Hybrid Systems Control–Theory, Tools, Applications*, pages 31–55. 2009.

[20] A. Donzé and O. Maler. Systematic simulation using sensitivity analysis. In *Proceedings of HSCC 2007*, 2007.

[21] G. et al. Frehse. Spaceex: Scalable verification of hybrid systems. In *Proceedings of CAV 2011*, volume 6806 of *LNCS*, 2011.

[22] Sicun Gao, Soonho Kong, and Edmund M. Clarke. dReal: An SMT solver for nonlinear theories over the reals. In *Automated Deduction?CADE-24*, pages 208–214. Springer, 2013.

[23] Nicolas Gast and Bruno Gaujal. Markov chains with discontinuous drifts have differential inclusion limits. *Performance Evaluation*, 69(12):623–642, 2012.

[24] 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*, number 3927 in LNCS, 2006.

[25] Michał Kisielewicz. *Stochastic differential inclusions and applications*. Springer, 2013.

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

[27] T.G. Kurtz. Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes. *Journal of Applied Probability*, 7:49–58, 1970.

[28] A. B. Kurzhanski and P. Varaiya. On ellipsoidal techniques for reachability analysis. *Optimization methods and software*, 17(2), 2002.

[29] Oded Maler. Computing reachable sets: an introduction. Technical report, Tech. rep. French National Center of Scientific Research. www-verimag. imag. fr/maler/Papers/reach-intro. pdf, 2008.

[30] 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*, 9(67):376–388, 2011.

[31] 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.*, page 23522364, 2009.

[32] N. Ramdani and N. S. Nedialkov. Computing reachable sets for uncertain nonlinear hybrid systems using interval constraint-propagation techniques. *Nonlinear Analysis: Hybrid Systems*, 5(2):149–162, 2011.

[33] Grégory Roth and William H Sandholm. Stochastic approximations with constant step size and differential inclusions. *SIAM Journal on Control and Optimization*, 51(1):525–555, 2013.

[34] Damjan Škulj. Discrete time markov chains with interval probabilities. *International journal of approximate reasoning*, 50(8):1314–1329, 2009.

[35] Yannis C Stamatiou, Paul G Spirakis, Theodoros Komninos, and Georgios Vavitsas. *Computer Network Epidemics: Models and Techniques for Invasion and Defense.* CRC Press, Inc., 2012.

[36] Emanuel Todorov. Optimal control theory. *Bayesian brain: probabilistic approaches to neural coding*, pages 269–298, 2006.

[37] Max Tschaikowski and Mirco Tribastone. Approximate reduction of heterogenous nonlinear models with differential hulls. *submitted to a journal*, 2015.

[38] Piotr Zgliczynski and TA Kapela. Lohner algorithm for perturbation of odes and differential inclusions. *Discrete Contin. Dyn. Syst. Ser. B*, 11(2):365–385, 2009.