



## Original Research Article

# Semiclassical approximations of stochastic epidemiological processes towards parameter estimation using as prime example the SIS system with import



Luís Mateus<sup>a,\*</sup>, Peyman Ghaffari<sup>a</sup>, Urszula Skwara<sup>a,c</sup>, Filipe Rocha<sup>a</sup>, Maíra Aguiar<sup>a</sup>, Davide Masoero<sup>b</sup>, Nico Stollenwerk<sup>a</sup>

<sup>a</sup> Centro de Matemática e Aplicações Fundamentais, Universidade de Lisboa, Portugal

<sup>b</sup> Grupo de Física-Matemática, Universidade de Lisboa, Portugal

<sup>c</sup> Maria Curie-Skłodowska University, Lublin, Poland

## ARTICLE INFO

## Article history:

Received 29 January 2015

Received in revised form 4 September 2015

Accepted 5 September 2015

Available online 12 October 2015

## Keywords:

Stochastic processes

Master equation

WKB approximation

Eikonal approximation

Likelihood function

Kolmogorov forward equations

## ABSTRACT

In this paper we investigate several schemes to approximate the stationary distribution of the stochastic SIS system with import. We begin by presenting the model and analytically computing its stationary distribution. We then approximate this distribution using Kramers–Moyal approximation, van Kampen's system size expansion, and a semiclassical scheme, also called WKB or eikonal approximation depending on its different applications in physics. For the semiclassical scheme, done in the context of the Hamilton–Jacobi formalism, two approaches are taken. In the first approach we assume a semiclassical ansatz for the generating function, while in the second the solution of the master equation is approximated directly. The different schemes are compared and the semiclassical approximation, which performs better, is then used to analyse the time dependent solution of stochastic systems for which no analytical expression is known. Stochastic epidemiological models are studied in order to investigate how far such semiclassical approximations can be used for parameter estimation.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

To understand epidemiological processes it is important not only to model the phenomena but also to test the models against real empirical data (Stollenwerk et al. 2012). Simple statistical tools are often not helpful, since even the most basic epidemiological models are nonlinear, either in their deterministic ordinary differential equation formulation or, in the case of stochastic models, in their transition rates (Stollenwerk and Jansen, 2011). Stochastic models, formulated as time-continuous Markov processes, are necessary to study any kind of comparison with empirical data, which are noisy by nature. The dynamics of such stochastic processes are given by Kolmogorov forward equations, also called master equations in physics (Gardiner, 2004; van Kampen, 1992; Honerkamp, 1993).

Parameter estimation algorithms, in particular, need a large number of stochastic realizations or a good approximation of probabilities to describe the likelihood function for a given data set. The more efficient any approximation scheme is, the larger search in parameter space can be achieved to match the empirical data under investigation. Several schemes have been proposed to approximate the distributions of stochastic processes. Among these, the so-called semiclassical approximation seems to perform better than some of the other methods in many areas of parameter space.

In this paper we give a closer look to some of the recently suggested schemes, including semiclassical approximations (Maslov and Fedoryuk, 1981; Gang, 1987). The term, which refers to classical approximations of quantum-physical systems, means here that stochastic systems in their partial differential equation formulation can be approximated via simpler differential equation systems. This is of quite some interest for estimating parameters and deciding between competing models for the description of empirical systems and its data (Stollenwerk et al. 2012; Mateus et al., 2013). The semiclassical approximation can

\* Corresponding author. Tel.: +351 963108649.

E-mail address: [luisgam1@yahoo.com](mailto:luisgam1@yahoo.com) (L. Mateus).

also be used to study problems like, e.g., optimal paths to extinction (Lindley et al., 2014; Billings et al., 2013; Schwartz et al., 2011).

We begin by presenting the SIS system with import and explicitly calculating its stationary distribution. The Kramers–Moyal approximation towards a Fokker–Planck equation and its extension to a small noise approximation, also known as van Kampen system size approximation, are then used to approximate this stationary distribution analytically. These approximations compete with each other in modern parameter estimation algorithms, like iterated particle filtering, which can be applied to complex systems up to stochastic processes involving phenomena like deterministic chaos (Stollenwerk et al. 2012). Iterated particle filtering methods are essentially extensions of earlier approaches of comparing empirical data with simulations via  $\eta$ -ball neighbourhoods to obtain numerical likelihood functions (Stollenwerk et al. 2012; Stollenwerk and Briggs, 2000).

Two different approaches are described for the semiclassical approximation, via generating functions, as proposed in Elgart and Kamenev (2004), and a more direct method via approximation of the probabilities in the master equation directly, as initially suggested in Gang (1987) and also recently used in Billings et al. (2013). The second method has the advantage of proceeding directly from probabilities to the ansatz of an action function and its dynamics in a Hamilton–Jacobi equation, whereas the first method gives, via the extra step of the generating function, a Hamiltonian in simpler polynomials, which are quite common in theoretical physics. In the end, both ways give similar results, since they are linked by relatively simple transformations, as was already spotted in Gang (1987). We first describe the semiclassical approximation via the generating function for an SIS system with import, from which also dynamic solutions can be obtained for the first few moments, and then investigate the formulation via a direct approximation of the probability. The semiclassical results are compared with the ones obtained via the other approximation methods. We focus on the conceptual aspects of the semiclassical approximation, leaving to a future study the detailed evaluation of advantages and disadvantages of the different approximation schemes, especially in terms of computational efficiency.

Finally, after investigating the quality of the various approximation schemes, we analyse the time dependent distributions of stochastic systems for which no analytical expression is known. These time dependent solutions evolving from an exactly known initial state are the transition probabilities needed to construct the likelihood functions which are fundamental for parameter estimation.

The methods used in this paper for the stochastic SIS model with import can be generalized to higher dimensional systems. A brief generalization to the SIR system is done in Appendix B.

## 2. The stochastic SIS model with import

One of the simplest epidemiological models is the SIS system for susceptible and infected hosts, specified by the following reaction scheme



For stochastic variables  $I$  and  $S = N - I$ , with  $N$  the total population size of hosts, the model is a continuous-time Markov process with state space  $\{0, 1, \dots, N\}$ . The transitions are infection rate  $\beta$  and recovery rate  $\alpha$ . From the reaction scheme we can give the

dynamics of the probability  $p(I, t)$  in the form of a Kolmogorov forward equation or master equation (van Kampen, 1992; Honerkamp, 1993) as follows:

$$\begin{aligned} \frac{d}{dt} p(I, t) = & \frac{\beta}{N} (I - 1)(N - (I - 1)) p(I - 1, t) + \alpha(I + 1) p(I + 1, t) \\ & - \left( \frac{\beta}{N} I(N - I) + \alpha I \right) p(I, t). \end{aligned} \quad (2)$$

Given an initial condition  $p(I, t_0) = \delta_{I, I_0}$ , we have  $p(I, t) = p(I, t | I_0, t_0)$  and the master equation can be used to determine both the transition and the transient densities. With the definition of mean values, here  $\langle I \rangle := \sum_{I=0}^N I \cdot p(I, t)$ , we can derive the dynamics of the mean number of infected in time by inserting the above given dynamics of the probability and obtain, after some calculation (for a more extended derivation and further details see Stollenwerk and Jansen (2011), and additional literature references therein),

$$\frac{d}{dt} \langle I \rangle = (\beta - \alpha) \langle I \rangle - \frac{\beta}{N} \langle I^2 \rangle, \quad (3)$$

which now includes a higher moment  $\langle I^2 \rangle$ . For further investigation on moment closures see Stollenwerk et al. (2007).

To obtain a closed ordinary differential equation (ODE), we neglect the variance, restricting the dynamics to its simplest deterministic part, the so-called mean field approximation,  $\text{var} : = \langle I^2 \rangle - \langle I \rangle^2 \approx 0$ , and find the famous logistic equation

$$\frac{d}{dt} \langle I \rangle = \frac{\beta}{N} \langle I \rangle (N - \langle I \rangle) - \alpha \langle I \rangle \quad (4)$$

in a slightly unusual form, but mathematically equivalent to the well known formulation by Verhulst.

### 2.1. The SIS model with import

We can also consider in the SIS system an external import of infection, namely susceptibles inside our study population of  $N$  individuals become infected by meeting infected outside the study population, with ratio  $\varrho$  on top of the infection rate  $\beta$ . Or equivalently, infected from the outside visit the study area, but are not counted to belong to the study population  $N$ , and infect susceptibles from the study population. Then they contribute to the infected  $I$  from the study population with proportion  $\varrho N$ , hence in the infection term we have  $(I + \varrho N)$  to meet the  $S$  susceptibles, and infect them. The parameter  $\varrho$  is an import ratio per population size  $N$ . We have in the reaction scheme the following transitions



hence a direct transition from susceptibles in the study population to infected. The master equation is then given by

$$\begin{aligned} \frac{d}{dt} p(I, t) = & \frac{\beta}{N} (I - 1 + \varrho N)(N - (I - 1)) p(I - 1, t) + \alpha(I + 1) p(I + 1, t) \\ & - \left( \frac{\beta}{N} (I + \varrho N)(N - I) + \alpha I \right) p(I, t), \end{aligned} \quad (6)$$

from which all calculations can now be made analogously to the SIS system considered above.

The introduction of the import term has the advantage of being more realistic in empirical studies of infection at country or province level. Also, the master equation has a non-trivial stationary state, which is not the case for the SIS system without import, where  $p^*(I = 0) = 1$  and  $p^*(I) = 0$  for all  $I \in \{1, 2, \dots, N\}$  is the only stationary state.

Download English Version:

<https://daneshyari.com/en/article/4372347>

Download Persian Version:

<https://daneshyari.com/article/4372347>

[Daneshyari.com](https://daneshyari.com)