SULTAN MOULAY SLIMANE UNIVERSITY

DOCTORAL THESIS

Practical Estimation of Multidimensional Stochastic Differential Mixed Effects Model with Application to Insulin-Glucose Dynamics

---

Author:
Fadwa BAKRIM

Supervisor:
Prof. Hamid EL MAROUFY
Co-director:
Prof. Hassan AIT MOUSSE

Laboratory of Mathematics and Applications
&
Experimental Oncology and Natural Substances Team: Cellular and Molecular Immuno-Pharmacology
Department of Mathematics

July 28, 2021

SULTAN MOULAY SLIMANE UNIVERSITY

Abstract

Faculty of Sciences and Techniques

Department of Mathematics

Practical Estimation of Multidimensional Stochastic Differential Mixed Effects Model with Application to Insulin-Glucose Dynamics

by Fadwa BAKRIM

In this thesis, we consider a nonlinear and multidimensional stochastic differential mixed effects model, which enables the simultaneous representation of variability between experiments using mixed-effects approach, and the stochasticity in real processes dynamics using stochastic differential equations; considered as stochastic entities and containing random effects. Nevertheless, coupled stochastic differential equations with random effects often inherit a poorly posed estimation problem, since transition densities of most real processes are unknown, making all statistical approaches difficult and computationally expensive. Here, we propose a flexible and straightforward modelling framework to estimate the transition density using Fokker Planck equation. Finally, simulations from the two-dimensional Ornstein–Uhlenbeck (OU) process are addressed to evaluate the proposed methodology. Then, we assess the performance of our approach to estimate the stochastic minimal model with three coupled SDEs describing insulin-glucose kinetics, Such a stochastic approach for glucose-insulin modelling has to our knowledge not yet been studied in literature.

Acknowledgements

First and foremost, I would like to thank and express my deepest gratitude to my supervisor, Prof. Hamid El Maroufy, for the patient guidance, encouragement and advice he has provided throughout my doctorate study. His precious guidance helped me during all the time of research and writing of this thesis.

My sincere thanks also go to Dr. Taib Ziad, for his contribution to this research enlightening certain parts of this research for me, and for his insightful comments and encouragement.

I thank my fellow labmates for the stimulating discussions, and for the time we have worked together, especially Dr. Abdellah Abou-Bakre who has always been helpful.

Last but not the least, my thanks go to family and friends especially to those who have had a trace in the pursuit of this research.

Contents

List of Figures

List of Tables

Symbols & Abbreviations

Symbols

Ω	the universal set or space or sample space
$\mathcal{F}$	σ-algebra on Ω
$(\Omega, \mathcal{F})$	measurable space
$\mathbb{P}$	Probability measure on $(\Omega, \mathcal{F})$
$(\Omega, \mathcal{F}, \mathbb{P})$	probability space
$\mathcal{F}_t$	Filtration
$\mathcal{C}^R(\mathbb{R})$	space of R continuous derivatives functions
$\mathbb{E}$	mathematical expectation
$\pi(.)$	unnormalized distribution
$\mathbf{p}(\theta)$	prior distribution of the parameter θ
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^d$	the set of ordered d-tuples $(x_1, x_2, \ldots, x_d)$
$\mathbb{N}$	the set of natural integers
$\mathcal{N}$	normal distribution
$\delta(A)$	the Dirac delta function $\delta(A)$ equal 1 if $0 \in A$ and 0 otherwise
$\mathcal{B}$	the Borel σ-algebra
$\mathbf{B}$	Brownian motion
$\sim$	Equivalent
$\propto$	Proportional

Abbreviations

WHO	**W**orld **H**ealth **O**rganization
SDEs	**S**tochastic **D**ifferential **E**quations
PK	**P**harmaco **K**inetic
PD	**P**harmaco**D**ynamic
MCMC	**M**arkov **C**hain **M**onte **C**arlo
PMCMC	**P**article **M**arkov **C**hain **M**onte **C**arlo
FP	**F**okker- **P**lanck
ML	**M**aximum **L**ikelihood
ODE	**O**rdinary **D**ifferential **E**quations
PDE	**P**artial **D**ifferential **E**quations
DDE	**D**elay **D**ifferential **E**quations
NLME	**N**on**L**inear **M**ixed **E**ffects

SDME	**S**tochastic **D**ifferential **M**ixed-**E**ffects
AD	**A**utomatic **D**ifferentiation
EM	**E**uler- **M**aruyama
EK	**E**xtended **K**alman
EKF	**E**xtended **K**alman **F**ilter
FOCE	**F**irst **O**rder **C**onditional **E**stimate
OU	**O**rnstein- **U**hlenbeck
AR-MH	**A**ccept-**R**eject **M**etropolis-**H**astings
GA	**G**enetic **A**lgorithm
EN	**E**lite **N**umber
SR	**S**election **R**ate
CP	**C**rossover **P**robability
MP	**M**utation **P**robability
SMM	**S**tochastic **M**inimal **M**odel
IVGTT	**I**ntra **V**enous **G**lucose **T**olerance **T**est
T1D	**T**ype **1** **D**iabetes
T2D	**T**ype **2** **D**iabetes

Publications & Communications:

Publications

Chapitre: Fadwa Bakrim, Hamid El Maroufy, and Hassan Mousse (Feb. 2020). "Simulation and Parametric Inference of a Mixed Effects Model with Stochastic Differential Equations using the Fokker-Planck Equation Solution". In: DOI: 10.5772/intechopen.90751.

Article: F. Bakrim, H. El Maroufy. "A Review on Estimation Methods of Nonlinear Mixed Effects Model With Stochastic Differential Equations, Application to Three Dimensional Ornstein-Uhlenbeck Process". Moroccan Journal of Quantitative and Qualitative Research,
Mathematical and Experimental works Vol 1, No 1 (2019).

Communications

- A participation by "an exact estimate for the stochastic minimal model of glucose and insulin homeostasis using the population-based approach", during the 6th International Congress of the Moroccan Society of Applied Mathematics, Mathematics SM2A6'2019, November 07-09, 2019, Beni Mellal, Morocco.

- A participation by "Parameter Estimation in Stochastic Differential Mixed-Effects Models of insulin sensitivity" during the fifth International Congress of the Moroccan Society of Applied Mathematics SM2A organized by the University MOULAY ISMAIL in Meknes,

- To strengthen our research tools, we participated in a seminar on writing and scientific communication taking place on April 1, 2017 at the Faculty of Legal, Economic and Social Sciences of Mohammedia, and in a training on -Data analysis with R - June 09, 2018 at the Faculty of Economics and Social Sciences of Mohammedia,

- We thus participated in the International Conference of Applied Mathematics (ICAM) taking place in Taza organized by Sidi Mohamed Ben Abdellah University, Engineering Sciences Laboratory.

- A presentation of "Estimation of high dimensional Nonlinear Mixed Effects Model with SDE (Stochastic Differential Equations)" was made during the sixth JSI Engineering Science Day, held on April 28, 2018 at the Ain-Chok Faculty of Sciences , Casablanca,

- And a participation by a presentation of a poster on "Estimation of high dimensional Nonlinear Mixed Effects Model with SDE (Stochastic Differential Equations) with an application to epidemiology" during the Afro-Mediterranean conference on multidisciplinary research and applications, on 07 and July 08, 2018, at the IAV institute, Rabat.

Chapter 1

General Introduction

1.1 Motivation

Motivation for the Statistical Methodology Research

An *SDME model* is established from the SDEs with the incorporation of random effects and stochastic components driven by the *Weiner process*, which can be understood as an extension of an ordinary differential equation model. For a deterministic differential equation model, the solution is a deterministic function, while the solution of an SDE is a continuous stochastic process called the *diffusion process*, which is *a continuous time Markov process*. The behaviour of a diffusion process is governed by its transition density, that is in turn governed by the values of the parameters in the *SDME model*. So, the inference on these parameters constitutes a central statistical interest around the *SDME models*.

The nonlinear mixed effects model incorporating *random effects* has become an increasingly popular choice for modeling real processes, due to its inherent incorporation of uncertainty, allowing simultaneous representations of randomness in dynamics of real processes and variability between experimental units. Therefore, it provides a powerful modeling tool with immediate applications. However, statistical inference for *SDME models* is not straightforward, since a closed form solution to many *SDME models* used in practice is not known, except for a few cases, and then the transition density for most others is not known. Moreover, in *SDME models*, explicit estimating equations for the *ML estimators* can be found after solving the integral in the marginal likelihood function of the parameters given the random effects. However, in general it is not possible to solve analytically and explicitly this integral.

Motivation for the Applied Research

The applied case of the research in this thesis concerns the disease of *diabetes*, which is affecting more and more people around the world. The World Health Organization (WHO) estimates that, in 2019, diabetes affected more than 463 million people worldwide , and that this number is presumably to increase to 578 million by 2030 and to 700 million by 2045. In 2005, around 1.1 million people died of diabetes. After cancer, injury and cardiovascular disease, *diabetes* is the fifth leading cause of death when classified by cause of death, of which nearly 80% of deaths occur in low- and middle-income countries. WHO predicts that deaths from *diabetes* will increase by more than 50% over the next ten years without urgent action. In addition, *diabetes* and its complications is one of the epidemics having a significant impact on the economics of health systems and countries. Also, WHO estimates that over the next ten years, China will lose 558 billion dollars of the value of its national income due to heart disease, stroke and *diabetes* alone. Indeed, given the complexity of the disease, recourse to the tools of applied mathematics is highly needed for the intervention and control of this epidemic, like the case of any other epidemic. Thus, models of dynamic systems are one of the essential mathematical tools to be able to develop strategies, whether it is from cell biology to pathophysiology, including pharmacology, chemistry, physics and engineering, transplantation and patient management and health care. Here, we will focus on classes of mechanistic models based on physiology; which is the minimal model describing the key components of the system's functionality, and which is able to measure the crucial processes of glucose metabolism and insulin control.

Earlier in the literature, the glucose-insulin kinetics described by the minimal model has solely been modeled as a deterministic model with only measurement errors. However, we believe that the processes of glucose and insulin should be described by stochastic differential equations, because in the deterministic version, they may not comply to the actual processes that take place inside the body. Therefore, here, we propose a stochastic version of the minimal model by adding stochastic components to the processes. Furthermore, we combine both the glucose and insulin parts of the minimal model to get a unified model, which is more physiologically sound but also more complex and highly ill-posed estimation problem.

So, modeling the disease of diabetes using an SDE model incorporating random effects, could in fact allow a more succinct modeling of system dynamics and individual deviation, with the advantage of estimating the metabolic portrait for an entire population. Such a problem could be analysed using a stochastic differential mixed effects

model.

1.2 Problem Statement

In statistical inference, the classical estimation procedure consists of maximizing the likelihood of the sample. However, a closed-form expression of the likelihood function for a stoachastic differential mixed effects model is rarely available, since the transition densities of the process are rarely known. Hence, exact maximum likelihood (ML) estimation is generally unrealizable, which is a problem often inherited in coupled stochastic differential equations. The problem then is that estimating parameters in SDME models is not straightforward, except for a few simple cases. In addition, the more the dimension of the random effects increases, the more the difficulties increase, both for the calculation of the transition density and for the integral of the density given random effects in the marginal likelihood of the parameters.

In this thesis, we deal with the development of a generic estimation approach, precise and achievable based on the ML estimation, which can be implemented in the absence of a closed expression of the transition density. In addition, as the case study of this thesis concerns the metabolism of diabetes, coupling the dynamics of glucose and insulin processes is more physiologically sound but also more complex and insoluble on the statistical-computer level, in particular when using the population-based approach.

1.3 Applied mathematics and Pharmacokinetics

The contribution of applied mathematics to biology and medicine is now evident. Indeed, the mathematical models in medicine concern very diverse subjects, for example, they are widely used for the treatments of cancer, in order to integrate the biological complexity and to provide algorithmic tools to the physicians to optimize the effectiveness of the anticancer treatments, while limiting their toxic effects. More precisely, statistics are an essential tool to validate the results of research in biology and medicine, e.g. study of the genome, spatial structure of living molecules (DNA, proteins), study of ecosystems, population genetics, phylogenetics, epidemiology, theory of evolution,...etc, while making it possible to quantify the hypotheses previously assumed. This is also the case for the current global situation regarding the corona virus, we firmly believe that statistical tools will have great results to study the behaviour of this pandemic in several dimensions: medical, economic, social and political, and that

the mathematical modelling will bring great results to the management of this health crisis by providing answers about the virus behaviour and its effects on the human body.

Among mathematical modelling tools in natural science, we have the pharmacokinetic (PK) model, which is built to predict the ADME phase: absorption (A), distribution (D), metabolism (M) and excretion (E) of natural or synthetic chemicals in humans or animals. Thus, pharmacokinetics (PK) presents a set of techniques which make it possible to generate and analyse the time-concentration relationship following the administration of a drug. The analysis of these relationships makes it possible to quantify by parameters the different processes following the course of some injections in blood. In this context, four methods of data analysis are generally used: 1) non-compartment analysis, 2) individual compartment analysis, 3) population compartment analysis and 4) physiology-based pharmacokinetics.

The last three methods are based on a compartmental modeling approach, because the characterization of ADME phases is done through a compartmental mathematical model. For methods 2 and 3, these are simple models, with a generally low number of compartments. In addition, the PK models describe the organism in a more mechanistic way, where each organ can be represented by one or more compartments. The models can thus reach several hundred compartments. It is therefore necessary to have more diversified data to identify the parameters. Thus, the differential equations are then used to describe the dynamic relationships determined by the PK methods, where the parameters model are estimated using appropriate statistical models. For example, for the individual compartment analysis method, where the data are studied on each individual, a statistical model is then added to describe variations in the value of parameters in the whole population. So, in general, the need to determine both positional parameters (i.e. fixed effects) and dispersion parameters (i.e. random effects) necessitates the use of nonlinear mixed effects models, see the following paragraphs.

Here, we focus on the individual compartment analysis method which is a modelling technique widely used in biology. It has a lot of applications in pharmacokinetics, metabolism, epidemiology and population dynamics, and which we will use later to describe glucose-insulin kinetics in the human body after a glucose injection, in order to formulate a suitable model for the metabolism of diabetes. Thus, compartmental model is the most complete model (Holz and Fahr 2001; Aarons 2005), it aims to contribute to researches, development of new drugs, to the assessment of chemical substance's toxicity risks and to biology. It can be classified into mechanistic models, which relate its parameters to physiological processes, but does not necessarily reflect

all functional entities of the organism, it is also physiologically based on well-defined and structured compartments interconnected by blood, biochemical and other lymphatics flows. As mechanistic model, it can be subdivided into compartments; designated as mammillary, where there is a central compartment interacting with a number of peripheral compartments that surround it; or as catenary compartment model which consists of a chain of interconnected compartments.

Moreover, in all areas, and there are many, where we can assess materials, the basic element is the compartment. By compartment, we designate two types of abstraction:

- Consider a region of space limited by barriers and a physical magnitude which has a property of homogeneity in this region.

- Or a substance or a physical quantity, without precise localization.

the compartment can therefore be something fictitious. For example, it can be a drug present in the blood or in an organ, or in a given population of all individuals carrying a particular pathogen. In the case of a drug administered orally, a distinction can be made, depending on the case, between the stomach, intestine, blood, kidneys, etc.

1.3.1 Population approach

The population approach treats simultaneously a group of individuals instead of a single individual by sharing information with all the entities in the population. This approach makes it possible to estimate the metabolic portrait of the entire population instead of individual metabolic portraits, which can prove very useful for studying the spread of an epidemy between individuals, as well as it leads to have statistical informations for each individual. In the literature, both Vicini and Cobelli 2001; Agbaje et al. 2003 have treated the minimal model describing the diabetes disease using the population approach, but only the kinetic glucose was taken into account in the model.

1.4 NonLinear Mixed Effects model

In some study contexts of certain phenomena, the experiment requires data on an entire population and not only on a single individual to obtain complete information on the phenomenon, as well as several repeated measurements of a quantitative variable for each unit, in order to model correctly the progression and development of a disease or an economic or a financial aggregate. Thus, for each individual, many repeated

measurements are taken at different points of time. This type of sampling is required in various fields, especially in pharmacokinetic/pharmacodynamic (PK/PD) applications and in biomedical researches. Thus, this kind of modelling leads to describing the common side of the phenomenon in a whole population and the specificity of each individual in relation to this phenomenon, which is deduced from the observations taken on each individual. It allows, therefore, to model the global behaviour of a phenomenon for a group of units and also its dynamic side.

It is often reasonable to consider that responses follow the same model structure for all experimental units, however, model parameters vary randomly among individuals. Therefore, there is an increasing popularity and an extreme need for mixed effects models, where both random and fixed effects are incorporated into the model, in various research fields. So, the introduction of mixed effects model occurred in PK/PD modeling, where we introduce fixed effects as common parameters for the whole population, and random parameters to model the specific behaviour of each unit having the same overall behaviour as the population with individual variants. Thus, both variations within and between groups are modeled, leading to a more precise estimation of population parameters, which is recommended especially in PK/PD studies, leading to considerable savings in both resources than human or animal discomfort. The population-based approach is then necessary in PK/PD data, where data from several subjects are considered simultaneously because it allows simultaneous estimation between individuals which gives a more robust estimation of parameters that may vary between groups depending on an underlying distribution.

In the theory, the stochastic differential equations (SDE)s have proved to be more useful than deterministic differential equations (ODE)s to describe the dynamic side of real processes in, e.g., the PK/PD phenomenon, finance studies (See: Brandt and Santa-Clara 2002), and other processes in different fields (See: Lánskỳ, Lánská, and Weiss 2004; Andersen and Højbjerre 2005; Ditlevsen and De Gaetano 2005b; Picchini, Ditlevsen, and De Gaetano 2006a; Picchini, GAETANO, and Ditlevsen 2010; Ditlevsen et al. 2007; Overgaard et al. 2007). In Choi and Rempala 2011, some examples of the application of the SDEs in the biomedical field are treated by the author, as well as other examples in pharmacokinetic field are discussed in Sheiner and Beal 1980; Sheiner and Beal 1981; Donnet and Samson 2013.

However, statistical inference for such model is not an obvious procedure, especially when the exact transition density of the process does not exist or cannot be available or approximated in a closed form. Moreover, even when the exact density exists, other constraints may occur which will not allows to get the exact estimators; when, for

example, the integral of the likelihood function cannot be analytically solved or it is complicated to obtain the gradient terms analytically; and the more the dimension random effects increase the more the degree of difficulty increases also.

1.4.1 Random effects

The need to incorporate random effects in modeling is required more and more, particularly in PK/PD modelling, because of their ability to model total variation, splitting it into its within- and between-individual components; and where responses of repeated measurements follow the same model form for all experimental subjects. But model parameters vary randomly among individuals, which involves incorporating random effects into the model. In other words, the fixed effects are used to describe the common population behaviour, while the random effects are used to account for the population variation on parameters for each individual. In general, it is of interest to obtain not only individual parameters but also a quantitative description of the parameter distribution across a population.

Since the population approach is increasingly used to analyse the PK/PD data, the random effects are used to account for the population variation on parameters of each individual in the population, these deviations can be explained in some way by difference in covariate values among groups. Therefore, these individual deviations are more efficient and useful in PK/PD modeling where data from several subjects of a population are considered simultaneously, since, in general, the estimation of individual parameters only may be insufficient. So, it is useful to obtain a quantitative description of the distribution of parameters in a population. Otherwise, combining stochastic fluctuations with a population approach is quite appealing but raises inference challenges as explained and discussed in the following sections. In general, the random effects are often assumed to be (multi)normally distributed, but it could be any well-behaved density function.

In addition, random effects modelling is an important task in the estimation approaches of mixed effects models, linear or not. In the literature, we sometimes see that the estimation of random effects is not interesting, and that the analyst is more interested in fixed effects and their estimation as well as in the parameters of the distribution of the random effects across the population. This view of the characterization of random effects is rather narrow, because, to precisely estimate the fixed effects in a model, the random effects have to be properly accounted for. Moreover, as random effects are useful for quantifying the variability in a population, and although we are generally

more interested in estimates of the population mean for a parameter, sometimes we are equally interested in the variability of the parameter between subjects in a population. Indeed, to do any type of Monte Carlo simulation of a model, we need both the mean and the variance estimate of the parameter. However, it is not the variance or the standard error of the estimate of a parameter being discussed, but how much the value of that parameter varies from one individual to another because, indeed, such variability makes the parameter a random effect, as opposed to a fixed effect with which no variability is associated.

1.4.2 NonLinear Mixed Effects model

The mixed effects model is a statistical model that is used for modeling responses of individuals that have the same global behaviour with individual variations, containing both fixed and random effects. In this kind of models, all the responses follow a common known functional form that depends on unknown effects, some of them are fixed representing the fixed effects for all individuals in a given population, and the others are random in order to account for individual deviations, and that can be due to the difference in covariate values in a population, with an underlying distribution. In addition, the Nonlinear mixed effects (NLME) models are useful in describing and modelling a nonlinear relationship between a response variable and parameters. They give information about variation of parameter values between groups and allow parameter estimates to vary among groups. Therefore, using NLME models within the population approach is an important tool to describe variability within the population.

Moreover, this kind of models allows by their flexible covariance for nonconstant correlation among observations and unbalanced data, which makes them a good choice in PK/PD modeling where it is expected to have both variabilities within and between individuals (See: Searle and McCulloch 2001a; Pinheiro and Bates 2006). In biomedical researchs, repeated measurements taken on a series of individuals or experimental units play an important role, and it is often reasonable to assume that responses follow the same structure of model for all experimental units, while parameters vary randomly among individuals. In this thesis, the NLME model will be applied later to estimate insulin sensitivity and other key parameters for modeling diabetes (See section 4). Thus, a nonlinear mixed effects model can be defined as:

$$z_{ij} = f(\Phi_i, t_{ij}) + \epsilon_{ij}, i = 1..N, j = 1...n_i \tag{1.1}$$

where z_{ij} is the jth observation at the moment t_{ij} of ith individual, and $f(\cdot)$ is a nonlinear function describing the relationship between response variable y and the individual specific parameter vector Φ_i, N is the number of subjects (units) and n_i is the number of repeated measurements taken for each subject, ϵ_{ij} is the residue of this individual model assumed independently and identically distributed according to the normal distribution with mean zero and variance σ^2, and represents the noise term of the model. Thus, we define the population model representing parameters of each individual as the following:

$$\Phi_i = A_{ij}\beta + B_{ij}b_i \tag{1.2}$$

where β represents the fixed effects vector described by the matrix A_{ij}, and b_i is the random effects vector that vary across individuals, determined by the matrix B_{ij}, its components vary between individuals independently and identically according to a distribution $P(b_i/\Psi)$ depending on a population parameters Ψ, which is usually assumed to be multivariate normal distribution with parameter vector Ψ. So, from (1.1) and (1.2), as noted before, we conclude that this type of model leads to having a common model structure for all the subjects of the experiment, where the parameters of the model vary randomly among the individuals. The introduction of random effects is very interesting in several studies, it also has many advantages for so-called joint modeling, in which we can study certain phenomenon together with the variables included in the model as covariates including random effects (See: Mamontov 2008).

1.4.3 Dynamical NLME model

To model the dynamics of biological processes, financial and economic data, it is necessary to incorporate systems of deterministic differential equations based on ordinary differential equations (ODE), partial (PDE) or delay (DDE). In this case, the mixed effects model with dynamic system can be written in the following form under two submodels. The first represents the continuous state equation defining the dynamics of the system, and the second is the discrete measurement equation, which defines the relationship between the states of the system and the obtained measurements. So, the NLME model with ODEs can be described as follows:

$$dy_t = f(y_t, \Phi_i, t, u_t)dt \tag{1.3}$$
$$z_{ij} = g(y_{ij}, \Phi_{ij}, t_{ij}, u_{ij}) + e_{ij}$$

where y_t is the state of the model at time t, u_t represents optional inputs at time t and e_t is a Gaussian white noise measurement error, with mean zero and variance-covariance matrix depending on Φ_i, $e_{ij} \sim \mathcal{N}(0, S(\Phi_i))$. However, real processes in different areas cannot be derived from deterministic mechanisms and be completely modeled, because, they are exposed to influences that are ignored or difficult to model explicitly and which the deterministic systems do not take into account; ignoring this effect in the modeling may affect the estimation of parameters and the derived conclusions. So, there is an increasing need to extend the deterministic models to models including stochastic components. A natural extension of deterministic differential equations model is a system of stochastic differential equations (SDE)s, by incorporating stochastic processes to the driving system equations or by modelling relevant parameters as suitable stochastic processes.

Moreover, the extension of ODEs to the SDEs makes it possible to explain the differences between the observations and the predictions by two types of noise: dynamic noise, that enters through the dynamics of the system and that can result from its random fluctuations or from the shortages of model, and the measurements error which are added in the case of an indirectly observed process, which may be due to a test error or to the existence of a disturbance and represent the uncorrelated part of the residual variability. In the theory, there are a rich and developed resources for mixed effects models whether deterministic (See: Vonesh and Chinchilli 1996; Searle and McCulloch 2001b; Kuhn and Lavielle 2005; Guedj, Thiébaut, and Commenges 2007; Wang 2007) or stochastic, linear or nonlinear models. For many applications of stochastic NLME models in biomedical field, see: Picchini, Ditlevsen, and De Gaetano 2006a; Picchini, GAETANO, and Ditlevsen 2010; Ditlevsen et al. 2007, and for the pharmacokinetic applications, see: Sheiner and Beal 1980; Sheiner and Beal 1981; Donnet and Samson 2013.

1.4.4 Stochastic NLME model

Extending to stochastic version of NLME models is achieved by adding an additional Wiener noise component (See next section). In fact, this additional noise allows handling of autocorrelated residuals originating from natural variation or systematic model error. Thus, models defined through stochastic differential equations allow for the representation of random variability in dynamical systems. So, incorporating a random component in NLME models remains an important method of analysis (See: Allen 2007a) to get good estimates. The stochastic version of nonlinear model mixed effects model is defined as follows:

$$\begin{aligned} dy_t &= f(y_t, \Phi_i, t, u_t)dt + \sigma(t, u_t, \Phi_i)dW_t \\ z_{ij} &= g(y_{ij}, \Phi_{ij}, t_{ij}, u_{ij}) + e_{ij} \end{aligned} \tag{1.4}$$

where σ is the diffusion term, and W_t is the Weiner process term.

Therefore, for a considered phenomenon, this class of models enables the simultaneous representation of randomness in the dynamics of the process and variability between experimental subjects, the stochastic NLME model is then a good applied mathematics tool for a powerful modelling that can be used in PK modeling. The first papers encouraging the introduction of stochastic components in PK/PD were published by D'Argenio and Park 1997; Ramanathan 1999b; Ramanathan 1999a, where authors underline that both deterministic and stochastic components have contributions in PK/PD: e.g. drug concentrations in the blood follow determinable trends, but the exact concentration at any given time cannot be fully determined. Also, a stochastic one-compartment PK model was proposed with a variable elimination rate in Ramanathan 1999a and more sophisticated PK models have then been proposed with multiple compartments, nonlinear or time-inhomogeneous absorption or elimination (See for example: Ferrante, Bompadre, and Leone 2003; Tornøe, Jacobsen, and Madsen 2004; Ditlevsen and De Gaetano 2005c; Ditlevsen, Yip, and Holstein-Rathlou 2005; Picchini, Ditlevsen, and De Gaetano 2006b). Moreover, parameter estimation for the NLME model with stochastic differential equations has been highly tackled in the literature, also motivated by financial applications (See: Sørensen 2004). However, many suggested solutions were proposed, but require high frequency data which is not suited for PK/PD data where designs are usually sparse.

1.5 An overview of the thesis

In this thesis, we are interested in the statistical inference of the mixed effects model with stochastic differential equations. We thus consider the glucose-insulin kinetics to study the metabolism of diabetes using the population approach by considering a set of individuals simultaneously, based on the multidimensional, nonlinear and stochastic minimal model. For this reason, we consider the Risken approximation using Fokker-Planck equation, giving two approximate closed forms for the estimation of the transition density of the process. The proposed estimate approach is then addressed by simulation studies to demonstrate that the proposed method provides accurate estimates. So, the main is to assimilate the observed data by estimating the relevant parameters of the model. The rest of the thesis is organized as follows:

Chapter 2– Background. : This chapter is devoted to some necessary concepts and mathematical objects for a good understanding of this manuscript. Then, we focus on the basics of the problem by giving a brief introduction to stochastic differential equations and Itô diffusion processes. Furthermore, we give a brief introduction to statistical inference tools and MCMC sampling methods.

Chapter 3– Statistical inference for stochastic differential mixed effects model. : In this chapter, we present the formulation of the stochastic differential mixed effects model, then we propose a fast approximate maximum likelihood procedure for the computation of the estimation of random and non-random parameters, since in most cases the likelihood function is not available. In addition, a review on practical methods for estimating the nonlinear stochastic differential mixed effects model is given.

Chapter 4– Insulin sensitivity modelling. : In this chapter, we model the glucose-insulin kinetics by a three-dimensional stochastic process, through three differential stochastic equations established from the compartmental analysis based on the minimal model. The aim is to estimate the parameters of the model, since the transition density of the process is unknown, we then proceed to the Risken approximation described in Chapter 3 in order to have an approximate closed form of the likelihood function of the diabetes model in a closed form.

Chapter 5– Implementation issues and numerical applications. : Finally, in this chapter, the parameter estimation method proposed in Chapter 3 is evaluated using simulations from a standard model; it is the two-dimensional Ornstein-Uhlenbeck (OU), which is one of the few stochastic processes with an exact transition density; it is then

applied to the stochastic minimal model where the process transition density is approximated using the two proposed forms, one of which leads to an explicit likelihood function.

Section 5.2.3– Conclusions and future directions. : The main contributions of this thesis are reported, with the discussion of the limits of the approach from there, we conclude some points for further research such as: model hypotheses, measurement techniques, modeling, numerical methods, calculation software, Probability and statistics. Then, a perspective overview is briefly presented by the following for the formulation of an epedimic SDE model using the *forward Kolmogorov differential equations*.

Chapter 2

Background

Contents

In this section, we present briefly a list of preliminary concepts and mathematical objects necessary to a good understanding of this dissertation. Therefore, we give a short introduction to stochastic differential equations and Îto process, then we present shortly the Markov chain Monte Carlo methods with the necessary references and documents.

2.1 General Probability Theory

In this section, we present definitions and concepts necessary for the rest of our topic, that are related to probability theory and stochastic processes to define an Îto stochastic integral. For more details, the reader is referred to the book Øksendal 2003, Chap. 2.

Definitions and notions

Definition 2.1.1 (σ-algebra)**.** *If Ω is a given set, then a σ-algebra $\mathcal{F}$ on Ω is a family $\mathcal{F}$ of subsets of Ω with the following properties:*

i− $\emptyset \in \mathcal{F}$

ii− $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$, *where* $F^C = \Omega \setminus F$ *is the complement of* F *in* Ω

iii− $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A := \bigcup_{i=0}^{\infty} A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Definition 2.1.2 (Probability measure)**.** *A probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that:*

i− $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$

ii− $A_1, A_2, \ldots \in \mathcal{F}$ *and* $\{A_i\}_{i=0}^{\infty}$ *is disjoint (i.e.* $A_i \cap A_j = \emptyset$ *if* $i \neq j$*) then*

$$\mathbb{P}\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mathbb{P}\left(A_i\right).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space, and when $\mathcal{F}$ contains all subsets $\mathbb{P}$–null, it is called a complete probability space. Given any family $\mathcal{U}$ of subsets of Ω there is a smallest σ-algebra $\mathcal{F}^{\mathcal{U}} = \sigma(\mathcal{U})$ containing $\mathcal{U}$, namely:

$$\mathcal{F}^{\mathcal{U}} = \bigcap \{\mathcal{F}; \mathcal{F} \ \sigma\text{-algebra of } \Omega, \ \mathcal{U} \subset \mathcal{F}\}$$

.

where $\mathcal{F}^{\mathcal{U}}$ is the σ-algebra generated by $\mathcal{U}$, and it is called the *Borel σ-algebra* on Ω if $\mathcal{U}$ is the collection of all open subsets of a topological space Ω, and its elements are called Borel sets.

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space, then a function $\nu : \Omega \longrightarrow \mathbb{R}^n$ is called $\mathcal{F}$-measurable if

$$\nu^{-1}(U) := \{\omega \in \Omega; \nu(\omega) \in U\} \in \mathcal{F}$$

for all Borel sets $U \subset \mathbb{R}^n$. A *random variable* Y is an $\mathcal{F}$-measurable function $Y : \Omega \longrightarrow \mathbb{R}^n$

If $Y : \Omega \longrightarrow \mathbb{R}^n$ is any function, then the σ-algebra $\mathcal{F}^Y$ generated by Y is the smallest σ-algebra on Ω containing all the sets

$$Y^{-1}(U); \quad U \subset \mathbb{R}^n \text{ open.}$$

Definition 2.1.3 (Stochastic process)**.** *A stochastic or random process $(Y_t)_{t\in\mathbb{T}}$ can be defined as a parametrized collection (family) of random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that is indexed by non-empty time set $\mathbb{T} \subseteq \mathbb{R}_+$. Therefore, the stochastic process $\mathbf{Y}$ can be written as a function:*

$$\begin{aligned} \mathbf{Y} : \ & \mathbb{T} \times \Omega \longrightarrow \mathcal{Y} \\ & (t, \omega) \longmapsto Y(t, \omega) \end{aligned}$$

with state space $\mathcal{Y} \subseteq \mathbb{R}^d$, $d \geq 1$.

But generally, dependency on ω is considered in the notation $\mathbf{Y} = (Y_t)_{t\in\mathbb{T}}$. In addition, for each $t \in \mathbb{T}$ fixed we have a random variable

$$\omega \longrightarrow Y_t(\omega); \quad \omega \in \Omega.$$

and that a path of Y is considered by fixing $\omega \in \Omega$ defined by the following function

$$t \longrightarrow Y_t(\omega); \quad t \in \mathbb{T}.$$

Definition 2.1.4 (Filtration)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family $(\mathcal{F}_t)_{t\geq 0}$ of sub σ-algebras of $\mathcal{F}$. In other words, for each t, $\mathcal{F}_t$ is a σ-algebras included in $\mathcal{F}$ and if $s \leq t$, $\mathcal{F}_s \subset \mathcal{F}_t$. A probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ endowed with a filtration $(\mathcal{F}_t)_{t\geq 0}$ is called a filtered probability space.*

Definition 2.1.5. *A stochastic process $(Y_t)_{t\in\mathbb{T}}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if, for every $t \in \mathbb{T}$, the random variable Y_t is $\mathcal{F}_t$-measurable.*

A stochastic process $\mathbf{Y}$ is always adapted to its natural filtration $\mathcal{F}_t^Y = \sigma(Y_s, s \leq t)$, which is at the same time the smallest filtration to which $\mathbf{Y}$ is adapted. Then, one of the most important classes of stochastic processes is the *Martingale*

Definition 2.1.6 (Martingale)**.** *A stochastic process $(Y_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ (and with respect to $\mathbb{P}$) if:*

i− Y_t *is $\mathcal{F}_t$-measurable for all $t \geq 0$,*

ii− $\mathbb{E}[Y_t] < \infty$ *for all $t \geq 0$,*

iii− $\mathbb{E}[Y_t|\mathcal{F}_s] = M_s$ *for all* $s \leq t$.

2.2 Brownian motion

2.2.1 The Wiener noise - Brownian motion

The Wiener process, named in honour of Norbert Wiener [1], is one of the most famous continuous stochastic processes $\{W(t)\}_{(t\geq 0)}$, and it is a modern model that describes the Brownian motion. It's similar to a random walk with a time step go to 0 in a proper way and can be defined by the following three conditions:

i) $W(0) = 0$

ii) $\{W(t)\}_{(t\geq 0)}$ has independent increments, the random variables: $W_{t_1}; W_{t_2} - W_{t_1}; ...; W_{t_k} - W_{t_{k-1}}$ are independent for all $0 \leq t_1 \leq t_2 \leq ... \leq t_k$.

iii) $W(t+s) - W(s) \sim \mathcal{N}(0,t)$ for all $t > 0$

Thus, from i), ii) and iii) the Wiener process is also Gaussian with mean zero and variance proportional to the time, which implies that the Wiener process cannot be stationary because its variance depends on t.

In many fields such as: physics, chemistry, biology and others, the mathematical description of Brownian motion is of importance. It was named for the Scottish botanist Robert Brown, the first to study such fluctuations (1827), and describes the random movement of particles by impacting the surrounding medium in a fluid.

Definition 2.2.1 (Brownian Motion)**.** *A real-valued* $\mathcal{F}$*-adapted process* $\mathbf{B} = (B_t)_{t\geq 0}$ *is a Brownian motion if it satisfies the following conditions:*

i− $B_0 = c$ *almost surly for all* $c \in \mathbb{R}$ *fixed,*

ii− *All paths are almost surly continuous,*

iii− *All paths have independent and stationary increments,*

iv− $B_t \sim \mathcal{N}\left(0, \sigma^2 t\right)$ *for all* $t \geq 0$ *and constant volatility parameter* $\sigma \in \mathbb{R}_+$.

$(B_t)_{t\geq 0}$ is a martingale with respect to its natural filtration $\mathcal{F}_t$ generated by $\{B_s,\ s \leq t\}$ Øksendal 2003, Chap 3. From *iii*− in Definition 2.2.1, the increments of the Brownian motion are statistically independent on non-overlapping intervals, i.e $B_{t_1} - B_{t_0}, B_{t_2} -$

[1] Norbert Wiener (November 26, 1894 – March 18, 1964), was an American mathematician and philosopher.

$B_{t_1}, B_{t_3} - B_{t_2}, \ldots$, $0 \leq t_0 < t_1 < t_2 < \ldots$, are pairwise independent, which means that the probability distribution function for $B_{s+t} - B_s$ is fixed for all $s \in \mathbb{T}$ such that $s + t \in \mathbb{T}$. When $c = 0$ and $\sigma = 1$ we talk about the standard Brownian motion, and d-dimensional (standard) Brownian motion if its d components are mutually independent. The probability law induced by standard Brownian motion is thus called Wiener measure.

2.2.2 Stochastic Differential Equations

Definition

Generally, a stochastic differential equation (SDE) is a generalization of the notion of differential equation taking into account a white noise term and which solution is a stochastic process. Using the SDE allows for the representation of random variability in dynamical systems, which is becoming more and more important (e.g. see: Allen 2007b; Øksendal 2003) and constitutes a standard tool for modelling biological, financial, neuronal and population growth dynamics. Thus, for phenomena whose dynamics are affected by random noise as in physics, SDEs are an established tool for modelling. So, the introduction of stochastic components to deterministic models is an important tool of analysis (See: Aït-Sahalia 2002), and is more appropriate to model the intra-individual variations rather than ODEs. As noted before, dynamical biological processes are usually modeled by means of systems of ODEs which do not account for the noisy components of the system dynamics often present in biological systems. The cumulative effect on the actual state; which cannot be individually included in the model description of the system of a host of mechanisms (like hormonal oscillations, variations of the stress level, variable muscular activity etc.); is represented by what is called system error (or system noise). Noise in the differential equations describing the behaviour of the system requires an extension to the class of stochastic differential equation (SDE) models. For application areas including econometrics and finance (See: Aït-Sahalia 2009; Aït-Sahalia and Jacod 2012; Eraker and Wang 2015; Aït-Sahalia and Hurd 2016; Aït-Sahalia et al. 2017; Aït-Sahalia and Xiu 2017; Eraker and Wu 2017).

Ito integral and stochastic differential equations

Definition 2.2.2 (Itô process). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, then an Itô process $\mathbf{Y} = (Y_t)_{t\geq 0}$ is a process which satisfies:*

$$Y_t(\omega) = Y_s(\omega) + \int_s^t \boldsymbol{\mu}\left(u, Y_t\right) du + \int_s^t \boldsymbol{\sigma}\left(u, Y_t\right) dB_u, \tag{2.1}$$

for any $[s,t] \subseteq \mathbb{T}$, where the functions $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are jointly $\mathcal{B} \times \mathcal{F}$-measurable, $\mathcal{F}$-adapted and satisfies, the criteria

$$\int_s^t \left|\boldsymbol{\mu}\left(u, Y_t\right)\right| du < \infty,$$

$$\int_s^t \left(\boldsymbol{\sigma}\left(u, Y_t\right)\right)^2 du < \infty,$$

An Itô process is a stochastic process that can be, formally, written as:

$$dY_t = \boldsymbol{\mu}\left(t, Y_t\right) dt + \boldsymbol{\sigma}\left(t, Y_t\right) dB_t.$$

and using the definition of the Itô integral, a stochastic process $\mathbf{Y}$ can also be defined as a solution of the stochastic differential equation (2.2)

$$dY_t = \boldsymbol{\mu}\left(t, Y_t\right) dt + \boldsymbol{\sigma}\left(t, Y_t\right) dB_t, \tag{2.2}$$

if and only if $\mathbf{Y}$ satisfies the stochastic integral equation (2.3)

$$Y_t(\omega) = Y_s(\omega) + \int_s^t \boldsymbol{\mu}\left(u, Y_t\right) dt + \int_s^t \boldsymbol{\sigma}\left(u, Y_t\right) dB_u, \tag{2.3}$$

almost surely. Then $\mathbf{Y}$ is an Itô process and one can prove that it is Markovian.

Existence and uniqueness

The existence and uniqueness of solutions to stochastic differential equations (2.2) is guaranteed by the following theorem:

Theorem 2.2.1 (Existence and uniqueness theorem for stochastic differential equations)**.** *Let* $\mathbb{T}^{'} \subset \mathbb{T}$ *and* $\boldsymbol{\mu}(\cdot) : \mathbb{T}^{'} \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\boldsymbol{\sigma}(\cdot,\cdot) : \mathbb{T}^{'} \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$ *be measurable functions satisfying*

$$\|\boldsymbol{\mu}(t,x)\|^2 + \|\boldsymbol{\sigma}(t,x)\|^2 \leq C\left(1 + \|x\|^2\right), \text{ (Linear growth)} \tag{2.4}$$

$$\|\boldsymbol{\mu}(t,x) - \boldsymbol{\mu}(t,y)\| + \|\boldsymbol{\sigma}(t,x) - \boldsymbol{\sigma}(t,y)\| \leq D\|x-y\|, \quad \text{(Lipschitz continuity)} \tag{2.5}$$

for $x, y \in \mathbb{R}^n, t \in \mathbb{T}^{'}$ *and some constants* C *and* D.

Let Z *be a random variable which is independent of the* σ*-algebra* $\mathcal{F}_\infty = \sigma(B_s, s \geq 0)$ *and such that* $\mathbb{E}|Z|^2 < \infty$.

Then, the stochastic differential equation

$$dY_t = \boldsymbol{\mu}(t, Y_t)\, dt + \boldsymbol{\sigma}(t, Y_t)\, dB_t, \quad t \in \mathbb{T}^{'}, Y_0 = Z \tag{2.6}$$

has a unique t-continuous solution $Y_t(\omega)$ *with the property that* $Y_t(\omega)$ *is adapted to the filtration* $\mathcal{F}_t^Z = \sigma(Z, B_s, 0 \leq s \leq t)$ *and:*

$$\mathbb{E}\left[\int_{\mathbb{T}'} |\mathbf{Y}_t|^2\, dt\right] < \infty \tag{2.7}$$

The solution $\mathcal{Y}$ is $\mathcal{F}_t^Z$-adapted and it is called a strong solution. However, an explicit solution of an SDE is usually not found, but an explicit solution to a family of stochastic differential equations is well discussed in Kouritsin and Deli 2000.

2.3 Likelihood function and Markov process

Markov process, named for the Russian mathematician Andrey Andreyevich Markov (1856 – 1922), is one of the most important of all random processes. It defines as follows:

Definition 2.3.1 (Markov process)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and* $\mathbf{Y} = (Y_t)_{t\geq 0}$ *be a stochastic process defined on this space. The* σ*-algebra* $\mathcal{F}_t = \sigma(Y_s, 0 \leq s < t)$, $t \geq 0$, *presents process states on the past of the process and time* t.

The real-valued, $\mathcal{F}_t$*-adapted stochastic process* $\mathbf{Y}$ *is called a Markov process if the Markov property*

$$\mathbb{P}(Y_t|\mathcal{F}_s) = \mathbb{P}(Y_t|Y_s) \quad \textit{holds a.s. for all } 0 \leq s \leq t < \infty$$

Therefore, by knowing the past states of the process one can predict the probabilities related to the values of the states of its future, as well as by knowing all the past information of the process before the time s. The future states of the process depend on those of the past through the present because the process only knows Y_s and does not know how it got there, that is, once the present is known, the past and the future are independent.

Transition density

For an Itô process $\mathbf{Y}$ being a solution of the SDE (2.2), it verifies the Markov property above and its *transition density* $p(s, \mathbf{y}, t, \mathbf{z})$ is defined as:

$$p(s, \mathbf{y}; t, A) = \mathbb{P}(Y_t \in A | Y_s = y) = \int_A p(s, \mathbf{y}; t, \mathbf{z}) dz, \tag{2.8}$$

for all $\mathcal{F}$-measurable sets $A \subseteq \mathcal{Y}$. The density $p(s, \mathbf{y}; t, \mathbf{z})$ of a Markov process $\mathbf{Y}$, from state $\mathbf{y} \in \mathcal{Y}$ at time $s \geq 0$ to $\mathbf{z} \in \mathcal{Z}$ at time $t > s$, is defined as:

$$p(s, \mathbf{y}; t, \mathbf{z}) = \delta(\mathbf{y} - \mathbf{z})$$

where δ denotes the Dirac delta function. Moreover, when $\mathbf{Y}$ is homogeneous in time, the transition density depends on s and t only by their difference $t - s$, we also write $p(t - s; \mathbf{y}, \mathbf{z})$.

Likelihood function for discretely observed processes

The availability of the likelihood function in an explicit form is one of the most important issues in statistical inference, which requires the existence and knowledge of the form of the probability density function.

For a given realizations $y_1, y_2, \ldots, y_n$ of sample random variables $Y_1, \ldots, Y_n$ which are independents and identically distributed from a population with probability density function $p(y; \theta)$, the *likelihood function of the sample* has the form:

$$\mathcal{L}(\theta; y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} p(y_i; \theta). \tag{2.9}$$

For computational expedience, $\mathcal{L}$ will be transformed to the *logarithmic-likelihood function*

$$\ln \mathcal{L}(\theta; y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n} \ln p\,(y_i; \theta)\,. \tag{2.10}$$

the logarithmic likelihood function (2.10) is a function of θ which presents the probability of the observed random sample. To obtain an estimate of some unknown population parameters $\theta \in \Theta$, the method ML based upon the *principle of maximum likelihood* is applied. This method consists of maximizing the probability of observing the given random sample, by selecting the value of the parameter $\hat{\theta}$ as an estimate of θ . So, to find the estimates value of the parameters $\hat{\theta}$, we maximize (2.10) with respect to θ. The given value of *maximum likelihood estimate* $\hat{\theta} = \arg\max_{\theta} \mathcal{L}(\theta; y_1, y_2, \ldots, y_n)$ represents the most likely value of the parameter θ, to have generated the sample realizations $y_i,\ \ i = 1, \ldots, n$.

For a continuously observed diffusion processes on a finite interval, the statistical inference is based on the likelihood of the diffusion obtained using the Girsanov formula (See e.g. Liptser and Shiryaev 2001; Kutoyants 2004). For discretely observed diffusion processes, their likelihood depends on the transition densities of the diffusion $P(Y(t_k) \in A | Y(t_k - 1) = y)$, but in the most cases the likelihood is unavailable because these transition densities with respect to the parameters θ are not explicit and have no closed form (See for example, El Maroufy, Omari, and Taib 2012). So, since it is usually impossible to give a general formal solution of stochastic differential equations, this approach based on ML could be one of the serious challenges, as it involves the calculation of the transition density for this process which is often complicated or impossible. This problem is one of the most important questions to be addressed in this thesis.

2.4 Corresponding forward equation of Itô diffusion

The stochastic process has many fluctuations that its exact position cannot be determined but can be known for a region by its probability density; using the *Fokker–Planck* equation, such a probability density can be determined. The *Fokker–Planck* equation is a differential equation for the distribution function describing a Brownian motion by which the probability density of the stochastic process can be calculated in a much simpler way by solving this equation. This motion equation is usually used for variables describing a macroscopic but small system, where the fluctuations are important as

for some cases in physics, e.g., the position and the speed of the Brownian motion of a small particle. However, it can be also used for the larger system where, in spite of their small fluctuations, the stochastic description remains necessary when the deterministic equations may not be stable for this type of system.

Let $\mathbf{Y}$ be an Itô diffusion process :

$$dY_t = \boldsymbol{\mu}(t, Y_t)dt + \boldsymbol{\sigma}(t, Y_t)dB_t, \;\; \mathbf{Y}_{t_0} = y_0,$$

Under some regularity conditions on the coefficients $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ to ensure the existence and uniqueness. The *Fokker–Planck* equation associated with the Itô process is given by the following formula:

$$\frac{\partial p}{\partial t} = -\frac{\partial \left(p\boldsymbol{\mu}\right)}{\partial y} + \frac{1}{2}\frac{\partial^2 \left(p\boldsymbol{\sigma}^2\right)}{\partial y^2}.$$

Assuming that the derivatives of its transition density $p = p(t, y, z)$ in this partial differential equations exist and are continuous. Then, the equivalent *Kolmogorov forward equation* is:

$$\frac{\partial p(s, \boldsymbol{y}, t, \boldsymbol{z})}{\partial t} = -\sum_{i=1}^{d} \frac{\partial}{\partial \boldsymbol{y}_i}(\boldsymbol{\mu}_i(t, \boldsymbol{z})p(s, \boldsymbol{y}, t, \boldsymbol{z})) + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial \boldsymbol{y}_i \boldsymbol{y}_j}(\boldsymbol{\Sigma}_{i,j}(t, \boldsymbol{z})p(s, \boldsymbol{y}, t, \boldsymbol{z})), \tag{2.11}$$

for fixed $\boldsymbol{z}$ and t, where $\boldsymbol{y}, \boldsymbol{z} \in \mathbf{Y}$ and $t > s \geq 0$, and i, j denote the respective components of y, z, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$, where $\boldsymbol{\sigma}^T$ is the transposed vector of $\boldsymbol{\sigma}$.

We notice that the equation only determines the transition density p, and therefore the diffusion processes are already completely defined by their instantaneous mean and variance $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Moreover, the reverse implication of the property is valid, when the transition density of a stochastic process satisfies the Fokker-Plank equation (2.11), then it is an Itô diffusion process.

2.5 Markov chain Monte Carlo methods

Markov Chain Monte Carlo (MCMC) methods are a class of methods of sampling probability distribution functions or probability density functions (pdfs), based on the path

of Markov chains. It can be applied either in discrete or continue spaces sampling, these pdfs may then be either probability mass functions or probability densities. One of the advantages of MCMC methods, which makes them widely used, is that in probabilistic inferences a full analytic description of the properly normalized pdf is not required for sampling to proceed, We need only to compute ratios of the pdf at pairs of locations. This makes MCMC methods ideal when we want to sample the pdf for the parameters θ given the data x defining their *posterior pdfs*. The posterior pdf $p(\theta \mid x)$ is constructed from the pdf for the data given the parameters; presented by the likelihood $p(x \mid \theta)$; and by the prior pdf $\mathbf{p}(\theta)$ for the parameters using the famous Bayes formula known as "Bayes rule":

$$p(\theta \mid x) \quad = \quad \frac{1}{Z}\, p(x \mid \theta)\, p(\theta) \quad .$$

where Z is a constant often written as $\mathbf{p}(x)$ and known by the names "evidence", "marginal likelihood", "Bayes integral", and " prior predictive probability". Among the major problems encountered in this context, is that the likelihood (or the prior) is extremely hard to calculate, because it can have extremely complex structure, with multiple arbitrarily compact modes, arbitrarily positioned in a high dimensional parameter space θ. Therefore, the factor Z is often difficult to compute, and the function $p(\theta \mid x)$ is often up to a constant factor. So, this leads to compute ratios of the pdf at pairs of points but not the precise value at any individual point.

Moreover, one of the most important property of the MCMC method is the normalization-insensitive property. It means that it can be run without computing any derivatives or integrals of the function, in its simplest forms it is extremely easy to implement. For all these reasons, MCMC methods are still one of the very interesting statistical tools for scientists to sample posterior pdfs in real and complex situations where they find themselves. Here, we restrict our study to Gibbs Sampler and Metropolis-Hastings algorithms.

2.5.1 Gibbs Sampling Algorithm

This method is a Markov chain Monte Carlo method which allows us to sample a sequence of observations which are approximated from a specified multivariate probability distribution using only the conditional distributions (See: Casella and George 1992 for an explanation of Gibbs sampler with theory and examples, and Wilkinson

2006 for a review of Gibbs sampler). The reader can refer to Wilkinson 2006 for a simple application example of Gibbs sampling algorithm. Gibbs sampling was first used by Geman and Geman 1984 to study image processing models, before becoming a very popular MCMC method used in different fields (See: Carter and Kohn 1994; Gilks, Best, and Tan 1995; Arminger and Muthén 1998; Porteous et al. 2008; Damlen, Wakefield, and Walker 1999).

In Algorithm 1, we provide an algorithm for Gibbs sampler which generates samples of θ and x iteratively, from the conditional distributions $\pi(\theta \mid x)$ and $\pi(x \mid \theta, y)$, respectively.

Algorithm 1 Gibbs Sampler

1: Given an observed-data y.
2: Initialize x by sampling $x \sim \pi(. \mid y)$
3: **repeat**
4: Sample $\theta \sim \pi(\theta \mid x)$ using current x ▷ $\pi(\theta \mid x)$ is known analytically
5: Sample $x \sim \pi(x \mid y, \theta)$ using current θ
6: Store θ as a sample
7: **until** the desired number of samples for θ is reached

2.5.2 The Metropolis Hastings Algorithm

The Metropolis-Hastings algorithm is also one of the Markov chain Monte Carlo sampling methods, which is also widely used by scientists in many applications and one of the most needed statistical tools in situations difficult and complicated (See: Jeliazkov 2001; Geweke and Tanizaki 2001; Roberts and Stramer 2001; Cauchemez et al. 2004; Demiris and O'Neill 2005; Pratola 2016; Adaszewski et al. 2018). At first, it was proposed by Metropolis et al. 1953 in the field of statistical mechanics, and generalized later by Hastings 1970, the reader can refer to Hitchcock 2003; Chib and Greenberg 2012 for more descriptions of the Metropolis-Hastings sampling.

In contrast to Gibbs Sampler, the true conditional distribution $\pi(\theta|x)$ is not required in Metropolis-Hastings sampling, which makes this algorithm perfectly suited in many complicated cases, as it is generally not possible to sample from the true conditional distribution as is required for the Gibbs sampler. In Metropolis-Hastings algorithm, we consider a *proposal function* $q(.|\theta)$ from which we generate samples of θ iteratively and then accept the proposal sample with an acceptance probability. Given the posterior density $\pi(\theta|x)$ and a proposal density $q(.|\theta)$, then the Metropolis-Hastings is provided by the Algorithm 2 (For more details, see: Hastings 1970; Robert and Casella 2004).

Algorithm 2 Metropolis-Hastings algorithm Robert and Casella 2004

1: Given an observed-data x.
2: Initialize $\theta^{(0)}$
3: **for** $i = 1, \ldots, N$ **do**
4: Sample a candidate $\widetilde{\theta} \backsim q\left(\theta \mid \theta^{(i-1)}, x\right)$ (using a Gibbs sampler step given in Algorithm 1)
5: Calculate $\alpha = \min\left(1, \frac{\pi\left(\widetilde{\theta}|x\right)}{\pi\left(\theta^{(i-1)}|x\right)} \frac{q\left(\theta^{(i-1)} \mid \widetilde{\theta}|x\right)}{q\left(\widetilde{\theta} \mid \theta^{(i-1)}, x\right)}\right)$
6: Accept or reject $\widetilde{\theta}$ with probability α
7: Update $\theta^{(i)} = \begin{cases} \widetilde{\theta} & \text{with probability} \alpha \\ \theta^{(i-1)} & \text{else} \end{cases}$
8: Store $\theta^{(i)}$ as a sample
9: **end for**

2.6 Numerical methods for solving stochastic differential equations

Usually, a formal general solution of the Itô's stochastic differential equation (2.2) can not be available in a closed form. Therefore, using numerical methods to calculate approximations is extremely needed (See Panik 2017, Chap.7 for more knowledge about approximating methods). In this case, a discrete-time approximation is used to iteratively approximate a solution to (2.2), as in this method, the purpose is to discretize the data, using a recursive algorithm which produces the values in discrete time on a finite subinterval $[t_0, t] \subset \mathbb{T}$. While the approximation is made only at the discretization points, we will always view a discrete-time approximation as a "continuous-time process" defined on $[t_0, t]$. We deal in this thesis with two famous numerical methods which are widely used to approximate the process in (2.2), that is Euler-Maruyama (EM) and Milstein approximations (See Øksendal 2003; Elerian 1998 for more details on the description of the methods).

The EM Approximation

The EM approximation scheme is the simplest numerical approximation method used to approximate an Itô process for (2.2), based on a discrete-time recursive routine, given the time discretization $t_0 < t_1 < t_2 < \ldots < t_N = t$ of $[t_0, t]$, as follows:

$$\begin{aligned} Z_{i+1} &= Z_i + f\left(Z_i, t_i\right)\left(t_{i+1} - t_i\right) + g\left(Z_i, t_i\right)\left(B_{t_{i+1}} - B_{t_i}\right) \\ &= Zi + f\left(Z_i, t_i\right)\Delta_i + g\left(Z_i, t_i\right)\Delta B_i, \qquad i = 1, 2, \ldots, N-1, \end{aligned} \tag{2.12}$$

the given process from the EM approximation $Z = (Z_t)_{t\in[t_0,t]}$ is a continuous-time stochastic process, with $Z_{ti} \equiv Z_i$, $Z_0 = y_0$, $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, and $\Delta_i = t_{i+1} - t_i$, and when we have an equidistant discretization times $\Delta_i \equiv \Delta = (t - t_0)/N$, this time increments $\Delta_i, i = 0, 1, 2, \ldots$, should be "sufficiently small" to obtain a "good" approximate solution.

The Milstein Approximation

We shall now introduce the Milstein scheme which is an amelioration of the EM method by introducing a correction to the stochastic increment in (2.12), by introducing the term

$$\frac{1}{2} f(Z_i, t_i)\frac{\partial g}{\partial y}\left[\Delta {B_i}^2 - \Delta_i\right]$$

from the Itô–Taylor expansion, which gives the following scheme:

$$Z_{i+1} = Z_i + f(Z_i, t_i)\Delta_i + g(Z_i, t_i)\Delta B_i + \frac{1}{2} g(Z_i, t_i)\frac{\partial g}{\partial y}\left[\Delta {B_i}^2 - \Delta_i\right]. \tag{2.13}$$

The degree of precision (error) of the approximation can be measured by the expression $\mathbb{E}\left[Y_t - Z_t\right]$, since Y(t) and Z(t) are both random variables. For a continuous-time process **Y**, an approximation **Z** discretized in time converges of high order γ towards the solution **Y** at time t if:

$$\exists C \in \mathbb{R},\ \textit{such as}\ \mathbb{E}\left|Y_t - \mathbf{Z_t}\right| \leq C\Delta^{\gamma} \tag{2.14}$$

where Y_t is the true solution at time t with an approximate discretization Z_t, the constant C is not depending on Δ, and N is chosen large enough so that $\Delta = (t - t_0)/N\,(0, 1)$. On the other hand, the process **Z** converges weakly of order β to the solution **Y** at time t if a continuously differentiable polynomial function h and a constant C_h (independent of Δ) exist, such that:

$$|\mathbb{E}(Y_t) - \mathbb{E}(\mathbf{Z_t})| \leq C_h \Delta^{\beta} \tag{2.15}$$

Therefore, for the EM scheme the strong order of convergence is $\gamma = \frac{1}{2}$ (if f and g satisfy uniform growth and Lipschitz conditions) and converges with weak order $\beta = 1$, however, for Milstein scheme the strong order of convergence is $\gamma = 1$.

Chapter 3

Statistical inference for stochastic differential mixed effects model

Contents

3.1 Introduction

When both system noise and individual differences are considered, stochastic differential mixed effects (SDME) models ensue. This chapter is concerned with estimation methods for multidimensional and nonlinear dynamical models including stochastic differential equations and containing random effects (random parameters). This type of model has proved useful for describing continuous random processes, for distinguishing intra- and interindividual variability as well as for accounting for uncertainty in the dynamic model itself. Pharmacokinetic modeling, as seen before, often involves repeated measurements on a series of experimental units, and random effects are incorporated into the model to simulate the individual behaviour in the entire population. Unfortunately, the estimation of this kind of models could involves some difficulties, because in most cases, the transition density of the diffusion process given the random effects is not available. In this work, we focus on the approximation of the transition density of the stochastic process being solution of the SDEs model in a closed form, in order to obtain estimates for the model parameters using the Risken approximation using the corresponding forward equation of the process (See Section 3.5). Then, a simulation study is addressed later in Chapter 5 in order to provide results of the proposed methodology, with a real example of an SDME model applied in the epidemiology and based on the minimal model to describe glucose-insulin kinetics.

In the theory, there are rich and developed resources for mixed effects models whether deterministic (See: Vonesh and Chinchilli 1996; Searle and McCulloch 2001b; Kuhn and Lavielle 2005; Guedj, Thiébaut, and Commenges 2007; Wang 2007) or stochastic, linear or nonlinear. In this context, the reader is referred to see many applications of stochastic NLME models in biomedical fields in (Picchini, Ditlevsen, and De Gaetano 2006a; Picchini, GAETANO, and Ditlevsen 2010; Ditlevsen et al. 2007), and in pharmacokinetic studies in (Sheiner and Beal 1980; Sheiner and Beal 1981; Donnet and Samson 2013). Moreover, in Jelliffe, Schumitzky, and Van Guilder 2000, a review on methods for PK/PD population modeling is established, but the authors regret that system noise is not incorporated since it is difficult to estimate. Also, in Overgaard et al. 2005; Tornøe et al. 2005, a proposed SDE model with lognormal distributed random effects and a constant diffusion term is treated, but this constrains the class of models to be SDEs with additive noise. In Ditlevsen and De Gaetano 2005c, an example with the computation of the likelihood function in an explicit form was treated for a simple SDE model with random effects, however, generally the likelihood function is unavailable. Therefore, since the SDE models are more applied to biomedical data, there is

an increasing need for developing a general theory for parameter estimation of SDEs models incorporating random effects.

Parameter estimation in mixed effects models with SDEs, known by Stochastic Differential Mixed-Effects (SDME) models, is not an obvious procedure except in some simple cases (See: Ditlevsen and De Gaetano 2005a), because it is often difficult to write the likelihood function in its closed form. In the literature, the likelihood function of a nonlinear mixed effects model was approximated with the likelihood of a linear mixed-effects model (See Lindstrom and Bates 1990) . In this context, we propose a review on estimation methods of SDME models in Donnet and Samson 2008; Donnet and Samson 2013 and Bakrim and El Maroufy 2019, moreover, an example case that treats a generalized linear mixed models was proposed in Pinheiro and Chao 2006. In addition, to strengthen knowledge on estimation methods of SDME models, we refer to Overgaard et al. 2005; Tornøe et al. 2005 that propose an example of stochastic mixed effects model with random effects log-normally distributed with a constant diffusion term.

In general, it is difficult to obtain an explicit likelihood function because the transition density of the stochastic process is often unknown or that the integral in the marginal likelihood given the random effects cannot be computed analytically, and although the size of the random effects increases, the complexity of the problem increases also rapidly. Therefore, there is a significant need for approximation methods to compute the transition density in an approximate closed form, and also for efficient numerical integration methods to compute or approximate the integral in the likelihood function. For example, in the literature, the Laplacian and Gaussian quadrature approximation was widely used to approximate the integral in the likelihood (See: Searle and McCulloch 2001b; Picchini, GAETANO, and Ditlevsen 2010; Picchini and Ditlevsen 2011) as well as other numerical approaches (See: Fröberg 1985; Krommer and Ueberhuber 1998). Moreover, in the literature, several solutions have been proposed to approximate the transition density and have shown their effectiveness despite certain limitations. For example, the transition density could be approximated by *the solution of the partial differential equations of Kolmogorov* (See: Lo 1988); or by *the derivation of an Hermite expansion* of closed form at the transition density (See: Aït-Sahalia 2002; Aït-Sahalia 2002; Ait-Sahalia 2008), we notice that this method has been reviewed and applied for many known stochastic processes for one-dimensional and multi-dimensional time-homogeneous SDME model (See: Picchini, GAETANO, and Ditlevsen 2010; Picchini and Ditlevsen 2011); or by *simulating the process to Monte-Carlo-integrate the transition*

density (See: Nicolau 2002; Hurn, Lindsay, and Martin 2003; Ripley 2009). These techniques are very useful and can solve the problem, but unfortunately, they involve intense calculations which make the problem always complicated.

In this section, we focus on two fundamental issues concerning the implementation of SDEs in NLME models. The first is how the transition density of an SDME model can be approximated when it is not known, and the second is about approximating methods of the likelihood function when the integral given the random effects has no analytic solution. Then, we propose an optimization algorithm to obtain maximum likelihood estimators when the computation of gradients is so complicated or even impossible.

3.2 Formulation of stochastic differential mixed effects model

3.2.1 Itô formula

Itô's formula is one of the most important mathematical tools for stochastic calculus what the Newton-Leibnitz formula is for (the classical) calculus. It gives a practical method for the computation of stochastic integrals, and it is also used to relate differentiation and integration. Moreover, Itô's formula is very useful to evaluate Itô integrals and serve as a counterpart to the stochastic computation of the chain rule.

Theorem 3.2.1 (The general Itô formula)**.** *Let* $\mathbf{Y} = (Y_t,\ t \geq 0)$

$$dY_t = \boldsymbol{\mu}\left(t, Y_t\right) dt + \boldsymbol{\sigma}\left(t, Y_t\right) dB_t,$$

be a d-dimensional Itô process, as defined in Definition 2.2.2. Let $g : \mathbb{R}^d \times [0,1] \longrightarrow \mathbb{R}^l$ *be* $\mathcal{C}^2$ *map. Then the process* $\mathbf{Z}$ *such*

$$Z_t = g(Y_t, t)$$

is again an Itô process, and for its kth component we get the Itô formula, for $k = 1, \ldots, l$,

$$dZ_t(k) = \frac{\partial g^{(k)}(Y_t,t)}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial g^{(k)}(Y_t,t)}{\partial y^i} dY_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 g^{(k)}(Y_t,t)}{\partial y^i \partial y^j} dY_t^{(i)} dY_t^{(j)}, \qquad (3.1)$$

where

$$(dt)^2 = dt \cdot dB_t^{(i)} = dB_t^{(i)} \cdot dt = 0 \qquad \text{and} \qquad dB_t^{(i)} dB_t^{(j)} = \delta_{ij} dt,$$

and the upper indices denote the respective component numbers.

where δ_{ij} is the Kronecker delta, in combination with the SDE defining $\mathbf{Y}$. For more details and proof, (See Øksendal 2003, Chap4).

3.2.2 Stochastic Differential Mixed Effects model

Consider an N-dimensional continuous and stochastic process Y_t in the state space $E \subset \mathbb{R}^N$ described by the general first-order nonlinear stochastic differential equations of the Itô type (See: Ramanathan 1999c):

$$dY_t^i = \mu(Y_t^i, t, \theta, b^i)dt + \Sigma(Y_t^i, \theta, b^i)dW_t^i, \;\; Y_0^i = y_0^i, \;\; i = 1, ..., M, \tag{3.2}$$

where Y_t^i is defined as the solution of the SDME model (3.2) that exists under some conditions (See: Andersson and Britton 2012; Becker 1977; Oksendal 2003), and represents the observation of individual i from M different experimental units, $(i = 1, .., M)$, at the moment $t \geq t_0^i$, and $Y_0^i = Y_{t_0}^i$ is the initial state of Y_t for each subject. The process $\{(Y_t^i)_{t\geq 0}, i = 1..M\}$ is assumed to verify the same model structure (3.2) according to the individual deviations b^i; and $\theta \in \Theta \subset \mathbb{R}^p$ is a p-dimensional fixed effects vector which represents the same and common characteristics for all subjects; and $b^i \in B \subseteq \mathbb{R}^q$ are the q-dimensional individual random parameters assumed mutually independent, that vary between subjects according to a distribution of density $P_B(b^i|\Psi)$ depending on a population parameter Ψ; in the population approach, this parameter vector allows for a data from several subjects to be considered simultaneously. Each component b_l^i may follow a different distribution, $(l = 1, ..., q)$, and a standard choice for the joint density function $P_B(b^i|\Psi)$ of the vector b^i could be the Gaussian distribution; however, any other distributions may be considered continuous or discrete:

$$b^i \sim i.i.d\, \mathcal{N}(\vartheta, \phi).$$

The joint density function of the vector b^i is parameterized by a q-dimensional parameter $\vartheta \in \upsilon \subset \mathbb{R}^q$ and a $q \times q$-dimensional matrix $\phi \in \Phi \subset \mathbb{R}^{q\times q}$ representing the covariance matrix of b^i and specifying the parameters of the marginal distributions of the components b_l^i, $(1 \leq l \leq q)$; the components of ϕ and ϑ represent the population parameters Ψ. Moreover, we notice that, for Y_0^i, it is not necessarily to be known, and when its components are unknown, they must be considered as random effects since they often depend on measurements taken for each individual; however, in some cases

it can be known and assumed equal to a real constant. Also, we assume that the distribution of Y_t^i given (b^i, θ) and $Y_{t'}^i = y_{t'}$, $t' < t$, has a strictly positive density with regard to the Lebesgue measure on E:

$$y \to P_Y(y, t - t'|y_{t'}, b^i, \theta) > 0,\ y \in E. \tag{3.3}$$

$W^i(t)$ are the standard Brownian motions, and they are assumed mutually independent with b^j for all $1 \leq i, j \leq M$. The functions $\mu(\cdot) : E \times \mathbb{R} \times \Theta \times B \longrightarrow \mathbb{R}$ and $\Sigma(\cdot) : E \times \Theta \times B \longrightarrow \mathbb{R}^+$ represent, respectively, the drift and the diffusion term of the model and are assumed to have some properties sufficiently regular to ensure a unique solution to the model (See: Oksendal 2003).

The solution Y_t^i of (3.2) can be difficult to obtain in an explicit form, in this case, we approximate the different statistical characteristics of the process by Monte Carlo simulations that require discrete approximations of the continuous solution. In the theory, different schemes are available (See: Klöden and Platen 1992) with different levels of approximation quality and consumed-time. So, to approximate the solution Y_t^i in the time interval $[0; T]$, we consider the following time discretization:

$$0 \leq\ t_1\ \leq\ t_2\ \leq ... \leq\ t_j ... \leq\ t_N = T$$

and let $\triangle_j = t_{j+1} - t_j$ be the time step and $\triangle W_j = W(t_{j+1}) - W(t_j)$ the increments of the Wiener process with $\triangle W_j \sim \mathcal{N}(0, \triangle_j)$ which can be rewritten as: $\triangle W_j = \sqrt{\triangle_j} \cdot Z_j$ with $Z_j \sim \mathcal{N}(0, 1)$ for all j, which can be easily generated from the random normal number generator by the following code in Matlab software : dW = sqrt(dt)*randn.

According to the model in (3.2), the process Y is the same and follows the same structure for each individual in the population, the model describes the dynamic side of the individual behaviour following paths of Brownian motion. The individual deviations in the whole population is then modeled by both the different realizations of the Brownian motion paths $\{W_t^i\}_{t \geq t_0^i}$, and the incorporation of the random parameters b^i in the model. Therefore, the incorporation of parameters varying randomly between subjects allows to quantify the variability between individuals.

The goal, as it is explained before in the General Introduction, is to estimate the vector of fixed parameters θ and the parameter vector Ψ. However, the statistical inference for such models is a difficult issue, and the level of difficulty is not the same whether the

transition density is explicit or not and whether the process is observed directly or with measurement noise. In this work, we assume that the process was directly observed and no observation noise was considered.

3.3 Maximum likelihood estimation

The likelihood function of an SDME model is defined as follows:

$$L(\theta, \Psi) = \prod_{i=1}^{M} P(\underline{y}^i|\theta, \Psi) = \prod_{i=1}^{M} \int P_{\underline{Y}}(\underline{y}^i|b^i, \theta) P_B(b^i|\Psi) db^i \tag{3.4}$$

with :

$$P_{\underline{Y}}(\underline{y}^i|b^i, \theta) = \prod_{j=1}^{n_i} P_Y(y^i_j, \Delta^i_j|y^i_{j-1}, b^i, \theta), \tag{3.5}$$

where n_i is the number of observations for the subject i at discrete points of time $\{t^i_0, t^i_1, ..., t^i_{n_i}\}$, $i = 1, ..., M$ and $\Delta^i_j = t^i_j - t^i_{j-1}$, $j = 1, ..., n_i$. The conditional density $P_{\underline{Y}}(\underline{y}^i|\cdot)$ is equal to the product of the transition densities (3.5), for a given random effects b^i and θ, however, the availability of the transition density in an explicit form is rarely possible, which makes the statistical issues for the model (3.2) often complicated to obtain an exact likelihood function and exact ML estimators because computing the transition density is not always obvious and requires approximation methods. Nevertheless, there are some cases where the exact likelihood function is known and the exact ML estimators of θ are easily obtained (see references in the introduction). In fact, to compute the likelihood function in a closed-form for an SDME model, we can encounter two types of problems that require approximate methods to overcome them: First, when the transition density $P_Y(y^i_j, \Delta^i_j|y^i_{j-1}, b^i, \theta)$ is known but the integral in (3.4) has no solution, in this case, the numerical methods of approximation of the integral are required. Or, second, when $P_Y(y^i_j, \Delta^i_j|y^i_{j-1}, b^i, \theta)$ cannot even be expressed explicitly and must also be approximated (See next sections). Usually, in realistic examples, we have both an unknown transition density and an integral that is difficult to solve analytically. In theory, several methods for approximating transition densities and integrals have been proposed (See references cited in the introduction). Also, we notice

that, when the random effects have a discrete distributions; in that case the integral becomes a sum and can be easily computed when the transition density is known or apprpximated.

3.4 Closed approximate form of transition density and likelihood approximation

3.4.1 Likelihood approximation

When the transition density is not known, approximated solutions are required in order to get the ML estimators of the parameters. In the literature, many approximated methods to compute the transition density for the solution of an SDME model without measurement noise were proposed and have shown their effectiveness in spite of certain limitations. For example, Lo 1988 proposed to approximate the transition density by the solution of the partial differential equations of Kolmogorov, and **Ait2002a**, Aït-Sahalia 2002 and Ait-Sahalia 2008 proposed to approximate the transition density by the derivation of an Hermite expansion of closed-form. Also, in Pedersen 1995, Brandt and Santa-Clara 2002, Durham and Gallant 2002, Nicolau 2002, Hurn, Lindsay, and Martin 2003 and Ripley 2009 the transition density has been approximated by the simulation of the Monte Carlo process. Although these techniques have shown some usefulness in solving the problem, they unfortunately have drawbacks because they involve intense calculations, so that the problem is still not easy to solve.

Let $P_{\underline{Y}}^{(a)}(\underline{y}^i|b^i,\theta)$ be the approximation of (3.5), that we substitute in (3.4), where:

$$P_{\underline{Y}}^{(a)}(\underline{y}^i|b^i,\theta) = \prod_{j=1}^{N} P_Y^{(a)}(y_j^i, \Delta_j^i|y_{j-1}^i, b^i, \theta)$$

we obtain then an approximate likelihood function as follows:

$$L^{(a)}(\theta,\Psi) = \prod_{i=1}^{M} \int P_{\underline{Y}}^{(a)}(\underline{y}^i|b^i,\theta) P_B(b^i|\Psi) db^i \tag{3.6}$$

Usually, the integral in (3.6) have no closed solution, so the integration numerical methods are recommended. Therefore we cannot get the exact ML estimators $(\hat{\theta},\hat{\Psi})$, but,

by maximizing (3.6) over (θ, Ψ) we get approximate ML estimators $(\widehat{\theta}^{(a)}, \widehat{\Psi}^{(a)})$ that can show good properties.

In general, the classical inference of SDME models implies the problem of the numerical evaluation of the integral for the given random effects in the likelihood function, which becomes complicated especially when the model contains more than two random parameters. In the literature, several methods have been proposed and tested for the approximation of the integral, see references in the introduction and the following examples: see Fröberg 1985 and Krommer and Ueberhuber 1998 for the use of efficient numerical integration methods, and Picchini, GAETANO, and Ditlevsen 2010 for the Gaussian quadrature method for the case of SDME models with a single random effect, and Picchini and Ditlevsen 2011 for a general case with several random parameters using the Laplace approximation to compute the integral in (3.4) or (3.6) numerically. For the mixed effects framework, see Davidian and Giltinan 2003; Pinheiro and Bates 1995; Searle and McCulloch 2001b; Pinheiro and Chao 2006.

Gaussian quadrature approximation: It was proposed in Pinheiro and Chao 2006 and treated in Picchini, GAETANO, and Ditlevsen 2010 for the case of one random effect with normal distribution or any other continuous distribution (See Picchini, GAETANO, and Ditlevsen 2010 for more details). Assuming that $P_{\underline{Y}}(\underline{y}^i|b^i,\theta) \in \mathcal{C}^{2R}(\mathbb{R})$, the integral given random effect can be approximated by the Gauss-Hermite quadrature of an order (R) as the following:

$$\int P_{\underline{Y}}(\underline{y}^i|b^i,\theta)P_B(b^i|\Psi)db^i \simeq \sum_{r=1}^{R} \pi_r P_{\underline{Y}}(\underline{y}^i|\sqrt{2\phi}z_r + \vartheta, \theta) \tag{3.7}$$

where z_r, $r = 1, ..., R$ are the zeros of the Hermite polynomial $H_R(\cdot)$ of degree R and $\pi_r = \dfrac{2^{R-1}R!}{R^2(H_{R-1}(z_R))^2}$ are adequate weights, which does not depend on the individual. So, the likelihood (3.4) is approximated as follows:

$$L^{(R)}(\theta, \Psi) = \prod_{i=1}^{M} \sum_{r=1}^{R} \pi_r P_{\underline{Y}}(\underline{y}^i|\sqrt{2\phi}z_r + \vartheta, \theta) \tag{3.8}$$

We note that $L^{(R)}(\theta, \Psi)$ converge to the exact value of $L(\theta, \Psi)$ when the domain of integration is compact and $R \longrightarrow \infty$. Then, using optimization tools on (3.8) the approximated estimators $(\widehat{\theta}^{(R)}, \widehat{\Psi}^{(R)})$ are obtained :

$$(\widehat{\theta}^{(R)}, \widehat{\Psi}^{(R)}) = argmin_{(\theta,\Psi)}\{-\sum_{r=1}^{M} log(\sum_{r=1}^{R} \pi_r P_{\underline{Y}}(\underline{y}^i|\sqrt{2\phi}z_r + \vartheta, \theta))\} \tag{3.9}$$

Laplace approximation: For a multidimensional vector of random parameters, if the exact transition density or its closed-form approximation can exist, we can use the Laplace approximation method (See: Picchini and Ditlevsen 2011; Pinheiro and Chao 2006; Shun and McCullagh 1995), in order to obtain an explicit expression of the approximate likelihood function to maximize. So, for a q-dimensional random vector b^i, the likelihood function (3.4) can be approximated as:

$$logL(\theta, \Psi) \simeq \sum_{i=1}^{M} \left[logP_{\underline{Y}}(\underline{y}^i|\tilde{b}^i, \theta) + logP_B(\tilde{b}^i|\Psi) + \frac{q}{2}log(2\pi) - \frac{1}{2}log|-H(\tilde{b}^i|\theta, \Psi)| \right] \tag{3.10}$$

where:

$$\tilde{b}^i = argmax_{b^i}(f(b^i|\theta, \Psi)) \;\; and \;\; f(b^i|\theta, \Psi) = logP_{\underline{Y}}(\underline{y}^i|b^i, \theta) + logP_B(b^i|\Psi) \tag{3.11}$$

and $|\cdot|$ denotes the determinant of the Hessian matrix $H(b^i|\theta, \Psi)$:

$$H(\tilde{b}^i|\theta, \Psi) = \frac{\partial^2[logP_{\underline{Y}}(\underline{y}^i|\tilde{b}^i, \theta) + logP_B(\tilde{b}^i|\Psi)]}{\partial\tilde{b}^i\partial\tilde{b}^{iT}}$$

with $P_{\underline{Y}}(\underline{y}^i|b^i, \theta)$ is as in (3.5), then (3.10) is obtained by approximating $\int_B e^{log(f(b^i|\theta,\Psi))}db^i$ using a second-order Taylor series expansion, known as the Laplace approximation. In the case of using the Laplace method, the calculation of the Hessian matrix can be done analytically when it is possible, as for the examples in Chapter 5, or with help of a symbolic calculus software or the automatic differentiation tools (AD)(See: Griewank 2000).

Genetic Algorithm:

The genetic algorithm (GA) is a random search technique to look for an exact or approximated optimum points for optimization problems (See: Golberg 1989; Michalewicz 1992; Sivanandam and Deepa 2008). It is based on the concepts of natural genetic evolution that contains the following stages: reproduction, crossing, and mutation of a constantly evolving population. It sets up the evolution of a random population of potential solutions of N cardinal, then, the N simultaneous iterative trajectories interact with each other by following or imitating the biological evolution, for a convergence of some elements of the population towards an optimal point of the fitness function.

The GA can search in multiple directions to explore all the search space by the possibility of jumping across them, so that the seeds spread uniformly over the whole search space. In this algorithm, we have a diversity of initial populations which gives the global optimum faster than other algorithms, where the initial value is very important and should be enough close to the global optimum. All of these features allows the GA to be regarded as a driving tool of evolution giving good results for optimization processes (See: De Jong 2006; Michalewicz 1992). In the literature, there were many works on the application of GA in optimization problems as well as on the likelihood function (See: Yalçınkaya, Şenoğlu, and Yolcu 2018; Petrovski, Wilson, and McCall 1998). To generate the GA, we must first define some parameters of the algorithm: Population size N, EN, SR, CP, MP, fitness function, and convergence criteria. In the following we present the GA steps:

Steps of GA:

1. Generate initial population $\{\beta_1^{(0)}, \beta_2^{(0)}, ..., \beta_N^{(0)}\}, m = 0$ via an initialization strategy (random generation), in our case $\beta = (\theta, \Psi)$.

For $m = 0$:

2. Evaluate the Fitness function $log(-L^{(a)}(\theta^{(m)}, \Psi^{(m)}))$.

3. While (convergence criteria are not satisfied):

Do:

4. Replacement step (by using SR and EN): At the SR rate, individuals with the worst results in step 2 of Fitness function are replaced by new ones randomly generated , and a number EN of individuals is selected and accepted for the next step.

5. Selection Operator by using: Roulette Wheel Method, based on the fact that the more the individual has a good result of fitness function, the more likely he will be selected.

6. Crossover Operator by using CP & Mutation Operator by using MP: it is a mechanism of perturbation on the candidate individuals (parents) according to CP and MP to generate new groups of individuals and we obtain a new $(m+1)nd$ population $\{\beta_1^{(m+1)}, \beta_2^{(m+1)}, ..., \beta_N^{(m+1)}\}$.

Else:

7. Evolution stops, get GA output

8. $m = m + 1$

End For.

In this work, in Chapter 5, the GA is implemented using Matlab software, where the function "ga", to generate the Genetic algorithm, requires inputs that are chosen according to the constraints of each example (See the help window in Matlab). Furthermore, the algorithm parameters are chosen according to: De Jong 2006, as follows: $EN = 4$, $MP = 0.2$, $CP = 0.8$ *and* $SR = 1/3$, and the search spaces are around the confidence interval of the minimal model parameters (See: Andersen and Højbjerre 2005 and references therein).

3.5 An approach for a closed-form transition density

Here, we propose to approximate the transition density for a N-dimensional time-inhomogeneous SDME model (3.2) in a closed form using the Risken approximaton, the reader is referred to the book: Risken 1996, which is based on the *Fokker-Planck* (FP) equation characteristics or the forward Kolmogrov equation. The proposed methodology is then based on *the Kramers-Moyal expansion* that represents a motion equation verified by the probability density. Under some assumptions (See: Risken 1996), the probability density $\varphi(y,t)$ of a N-dimentional SDME model obeys *the Kramers-Moyal* (KM) expansion:

$$\frac{\partial \varphi(y_j,t_j)}{\partial t_j} = L_{KM}\varphi(y_j,t_j) \tag{3.12}$$

where:

$$L_{KM} = \sum_{n=1}^{\infty}(-\frac{\partial}{\partial y_j})^{(n)}T^{(n)}(y_j,\Delta_j|y_{j-1},b,\theta) \tag{3.13}$$

where the development coefficients are the moments, when the equation (3.13) stops after the second term, the obtained equation is then *the Fokker Planck* equation:

$$\frac{\partial \varphi(y_j, t_j)}{\partial t_j} = L_{FP}\varphi(y_j, t_j) \tag{3.14}$$

with L_{FP} is *the Fokker Planck* operator:

$$L_{FP} = -\frac{\partial}{\partial y_j}T^{(1)}(y_j, \Delta_j | y_{j-1}, b, \theta) + \frac{\partial^2}{\partial y_j^{(l)} y_j^{(k)}}T^{(2)}(y_j, \Delta_j | y_{j-1}, b, \theta) \tag{3.15}$$

where $T^{(1)}$ and $T^{(2)}$ are the drift and the diffusion term respectively, and l and k are the lth and kth componeneet of y respectively.

According to the special intitial condition: $\delta(y_1 - y_0) = P_Y(y_1, \Delta_1 | y_0, b, \theta)$, with the initial condition $\delta(y_j^i - y_{j-1}^i)$ is the Dirac-delta generalized function centered at y_{j-1}^i, the transiton density is the distribution $\varphi(y_j, t_j)$, thus, it must also obeys the equation (3.12). It represents therefore the solution of the motion equation. So, with respect to the special initial condition, the solution of *the Fokker Planck* equation (3.14) is the transition density for the model (3.2) (see: Risken 1996 and Lo 1988):

$$\frac{\partial P_Y(y_j, \Delta_j | y_{j-1}, b, \theta)}{\partial t_j} = L_{FP}P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) \tag{3.16}$$

with:

$$L_{FP} = -\frac{\partial}{\partial y_j^{(l)}}\mu_{(l)}(y_j, t_j, \theta, b) + \frac{\partial^2}{\partial y_j^{(l)} y_j^{(k)}}\Sigma_{lk}(y_j, t_j, \theta, b) \tag{3.17}$$

The equation (3.16) represents the motion equation of the process Y, and the resolution of this equation, when it is possible, leads to have an explicit form of the transition density. Then, for a small Δ_j, we get:

$$P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) = [1 + L_{FP}(y_j, t_j)\Delta_j + \circ(\Delta_j^2)]\delta(y_j - y_{j-1}) \tag{3.18}$$

So, bu inserting (3.17) in (3.18) then we may write up to terms of the order Δ_j^2:

$$P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) = [1 - \frac{\partial}{\partial y_j^{(l)}}\mu_{(l)}(y_j, t_j, \theta, b) + \frac{\partial^2}{\partial y_j^{(l)} y_j^{(k)}}\Sigma_{lk}(y_j, t_j, \theta, b)]\delta(y_j - y_{j-1}) \tag{3.19}$$

and by introducing δ in terms of Fourier integral, we get:

$$P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) = [1 - \frac{\partial}{\partial y_j}\mu(y_j, t_j, \theta, b)\Delta_j + \frac{\partial^2}{\partial y_j^{(l)} y_j^{(k)}}\Sigma(y_j, t_j, \theta, b)\Delta_j]\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iu(y_j - y_{j-1})}du$$

After a classical computation using the Gaussian integral properties such in Appendix A, and by replacing y_j by y_{j-1} in drift and diffusion terms, since: $\delta(y_j - y_{j-1})f(y_j) = \delta(y_j - y_{j-1})f(y_{j-1})$ we get the following:

$$\begin{aligned} P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) &= (2\sqrt{\Pi\Delta_j})^{-N}[Det\Sigma]^{-\frac{1}{2}} * exp(-\frac{1}{4\Delta_j}[\Sigma^{-1}]_{lk}[(y_j^i)^l - (y_{j-1}^i)_l - \\ &\mu_l(y_{j-1}^i, t, \theta, b^i)\Delta_j^i][(y_j^i)_k - (y_{j-1}^i)_k - \mu_k(y_{j-1}^i, t_{j-1}, \theta, b^i)\Delta_j^i]) \end{aligned} \tag{3.20}$$

In reality, the form (3.20) is not unique, we can deduce a class of equivalent forms, we can for example get the following expression that we will use in Chapter 5 (See the proof in the appendix) :

$$\begin{aligned} P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) &= (2\sqrt{\Pi\Delta_j})^{-N}[Det\Sigma]^{-\frac{1}{2}} exp(\frac{\partial}{\partial y_{j_l}}\mu_l(y_j, t_j, \theta, b)\Delta_j + \frac{\partial^2}{\partial y_{j_l}^2}\Sigma_l(y_j, t_j, \theta, b)\Delta_j \\ &-\frac{1}{4\Delta_j}[\Sigma(y_j, t_j, \theta, b)^{-1}]_{lk}[y_{j_l} - y_{j-1_l} - \mu(y_j, t_j, \theta, b)\Delta_j][y_{j_k} - y_{j-1_k} - \\ &\mu(y_j, t_j, \theta, b)\Delta_j]) \end{aligned} \tag{3.21}$$

The advantages of this approach, compared to those proposed in the literature for multidimensional SDME models with more than one random parameter Ait-Sahalia 2008, are that the computation of the approximate density is very easy and does not require a lot of time to compute in a software. Moreover, the proposed method is also effective even with large data, see the applied examples in Chapter 5. Also, the only task that can be time-consuming for the present methodology is in the optimization step to search for the optimum solution of the likelihood. The reader is referred to see other methods existing in the estimation methods literature in the following section of this chapter. Nevertheless, the method suffers some limitations, for example: when the conditions to use (3.20) or (3.21) are not verified when e.g. the inverse of the diffusion term does not exist or/and when the time step Δ_j is not sufficiently small, which require high frequency data and which is not suited for PK/PD data where designs are usually sparse.

Approximated estimators:

For a nonlinear SDME model with Gaussian random effects; using (3.6), (3.10) and (3.20); we obtain the following approximated likelihood function :

$$\begin{aligned} logL^{(a)}(\theta,\Psi) \simeq \sum_{i=1}^{M}[-\frac{q}{2}log((2\pi)) - \frac{1}{2}log(det(\phi)) - \frac{n_i}{2}log(det(\Sigma)) + (\sum_{j=1}^{n_i} log((2\sqrt{\Pi\Delta_j})^{-N}) \\ -\frac{1}{4\Delta_j}[\Sigma^{-1}]_{lk}[(Y_j^i)_l - (Y_{j-1}^i)_l - \mu_l(Y_{j-1}^i,t,\theta,\tilde{b}^i)\Delta_j^i][(Y_j^i)_k - (Y_{j-1}^i)_k - \mu_k(Y_{j-1}^i,t_{j-1},\theta,\tilde{b}^i)\Delta_j^i]) \\ -\frac{1}{2}(\tilde{b}^i-\nu)'\phi^{-1}(\tilde{b}^i-\nu) + \frac{q}{2}log(2\pi) - \frac{1}{2}log(det(-H(\tilde{b}^i|\theta,\Psi)))]. \end{aligned} \tag{3.22}$$

The ML estimators of (θ,Ψ) can be obtained using one of the optimization tools and numerical computation software, especially, when it is complicated to compute analytically the gradients of the likelihood. Here, we propose to use the genetic algorithm as an optimization tool to maximize the approximate likelihood function (3.22) using Matlab software:

$$(\hat{\theta},\hat{\Psi}) = argmin_{GA}(-log(L^{(a)}(\theta,\Psi))). \tag{3.23}$$

3.6 Practicale estimation methods of NonLinear stochastic differential Mixed Effects model

3.6.1 SDME model without measurements noise

The model in (3.2) is considered as an SDME model without measurement noise, where we assume that the process $(Y_t^i)_{t\geq t_0}$ is directly observed for all $i = 1...M$ at different time points t_{ij} for $j = 1...n_i$. This type of measurements represents the difference between two measurements taken simultaneously, which may be due to a test error or to existence of a disturbance element (for example a drug administered during a concentration measurement in plasma).

Estimation methods:

Hermite expansion of the approximated transition density: This method was originally proposed as an approximated method of transition density by **ait2002numerical** for unidimensional model with time-homogeneous equations, and was extended by Egorov, Li, and Xu 2003 to inhomogeneous equations, then to multidimensional equations by the original author Ait-Sahalia 2008. This method consists in transforming Y_t by the Lamperti transform ι, which exists when the diffusion is reducible and is defined as: $X_t \equiv \iota(Y_t) = \int_{Y_t} \frac{1}{\Sigma(u,\theta,b)} du$. The resulting process X_t is the solution of an SDME model with constant diffusion term equal to one and drift term defined as:

$$\mu_X^{(l)}(X_t) = \sum_{i=1}^{d} (\Sigma_{li}^{-1}(\iota^{-1}(X_t))\mu^{(i)}(\iota^{-1}(X_t))) - \frac{1}{2}\sum_{i,j,k}^{d} (\Sigma^{-1}(\iota^{-1}(X_t)) \frac{\partial \Sigma}{\partial Y_j}(\iota^{-1}(X_t))\Sigma^{-1}(\iota^{-1}(X_t)))_{li}\Sigma_{ik}(\iota^{-1}(X_t))\Sigma_{jk}(\iota^{-1}(X_t)). \tag{3.24}$$

Then we consider the transition density of the process X_t that is expanded in closed-form using an order $J = +\infty$ Hermite series, and approximated by a Taylor expansion up to order S. We assume that the functions $\mu(\cdot)$ and $\Sigma(\cdot)$ are infinitely differentiable in Y_t^i and three times continuously differentiable in θ and b^i, $\forall Y_t^i \in E$, $(\theta, b^i) \in \Theta \times B$; and that $\Sigma(\cdot)$ is bounded below by a strictly positive function. So, we obtain the following explicit sequence:

$$lnP_{\underline{Y}}^{(S)}(y_j^i, \triangle_j^i | y_{j-1}^i, b^i, \theta) = -\frac{d}{2} ln(2\pi \triangle_j^i) - \frac{1}{2} ln(det(\Sigma(y_j^i, \theta, b^i)\Sigma(y_j^i, \theta, b^i)^T)) + \frac{C_X^{(-1)}(\iota(y_j^i)|\iota(y_{j-1}^i))}{\triangle_j^i} + \sum_{s=0}^{S} C_X^{(s)}(\iota(y_j^i)|\iota(y_{j-1}^i)) \frac{\triangle_j^{s,i}}{(\,!s)} \tag{3.25}$$

where: $C_X^{(-1)}(x|x_0) = -\frac{1}{2}\sum_{h=1}^{d}(x^{(h)} - x_0^{(h)})^2$, $C_X^{(0)}(x|x_0) = \sum_{h=1}^{d}(x^{(h)} - x_0^{(h)}) \int_0^1 \mu_X^{(h)}(x_0 + u(x - x_0))du$, $C_X^{(k)}(x|x_0) = k\int_0^1 G_X^{(k)}(x_0 + u(x - x_0)|x_0)u^{k-1}du$, for $k = 1$ we have: $G_X^{(1)}(x|x_0) = -\sum_{h=1}^{d}\frac{\partial \mu_{X^{(h)}}(x)}{\partial x^{(h)}} - \sum_{h=1}^{d}\mu_{X^{(h)}}(x)\frac{\partial C_X^{(0)}(x|x_0)}{\partial x^{(h)}} + \frac{1}{2}\sum_{h=1}^{d}\left\{\frac{\partial^2 C_X^{(0)}(x|x_0)}{\partial x^{(h)2}} + \left(\frac{\partial C_X^{(0)}(x|x_0)}{\partial x^{(h)}}\right)^2\right\}$, and for $k \geq 2$ we have: $G_X^{(k)}(x|x_0) = -\sum_{h=1}^{d}\mu_{X^{(h)}}(x)\frac{\partial C_X^{(k-1)}(x|x_0)}{\partial x^{(h)}} + \frac{1}{2}\sum_{h=1}^{d}\frac{\partial^2 C_X^{(k-1)}(x|x_0)}{\partial x^{(h)2}} + \frac{1}{2}\sum_{h=1}^{d}\sum_{h'=0}^{k-1}\binom{k-1}{h'}\frac{\partial C_X^{(h')}(x|x_0)}{\partial x^{(h)}}\frac{\partial C_X^{(k-1-h')}(x|x_0)}{\partial x^{(h)}}$

Then, we obtain the following approximated log likelihood function of order S as follows:

$$logL^{(S)}(\theta, \Psi) = \sum_{i=1}^{M}\int logP_{\underline{Y}}^{(S)}(\underline{y}^i|b^i, \theta)P_B(b^i|\Psi)db^i \tag{3.26}$$

with $P_{\underline{Y}}^{(S)}(\underline{y}^i|b^i, \theta) = \prod_{j=1}^{n_i} P_Y^{(S)}(y_j^i, \Delta_j^i|y_{j-1}^i, b^i, \theta)$. So, the approximated estimators of parameters (θ, Ψ) are defined as:

$$(\widehat{\theta}^{(S)}, \widehat{\Psi}^{(S)}) = argmin_{(\theta,\Psi)}(-log(L^{(S)}(\theta, \Psi))) \tag{3.27}$$

Kolmogrov equation approximation: This method consists in approaching the transition density of a a continuous time stochastic process in an SDME model using Kolmogrov equations. It was was proposed by Lo 1988, who shows that under certain assumptions (See page: 6-8 in Lo 1988), the transition density satisfies the functional partial differential equation:

$$\frac{\partial}{\partial t}(P_Y(y_j^i, \triangle_j^i | y_{j-1}^i, b^i, \theta)) = -\frac{\partial}{\partial Y}[\mu(Y_j^i, t_j, \theta, b^i) P_Y(y_j^i, \triangle_j^i | y_{j-1}^i, b^i, \theta)] + \frac{1}{2}\frac{\partial^2}{\partial Y^2}[\Sigma(Y_j^i, \theta, b^i)^2 P_Y(y_j^i, \triangle_j^i | y_{j-1}^i, b^i, \theta)] \tag{3.28}$$

with the initial condition $P_Y(y_j^i, \triangle_j^i | y_{j-1}^i, b^i, \theta) = \delta(y_j^i - y_{j-1}^i)$ where $\delta(y_j^i - y_{j-1}^i)$ is the Dirac-delta generalized function centered at y_{j-1}^i, and with boundry condition: $P_Y(0, \triangle_j^i | y_{j-1}^i, b^i, \theta) = 0$. The cited reference has shown the utility of this approximation method by dealing with illustrated examples, and asserts the asymptotic properties of the obtained estimators. However, a limit of the efficiency of this method could be the possibility of being able to solve the particular partial differential equation of the process in question .

3.6.2 SDME model with measurement noise

In this paragraph, we consider that the process Y_t^i, $i = 1...M$, is not directly observed and we take into account the existence of a measurement error, we consider $Z_{t_j}^i = Z_j^i$ the jth observation of the individual i at the instant t_j, the whole of the model is then defined as:

$$\begin{aligned} &Z_j^i = h(Y_j^i, t_j, \theta, b^i) + g(Y_j^i, t_j, \theta, b^i)\varepsilon_{ij},\ \varepsilon_{ij} \sim \mathcal{N}(0, I_S), S < N \\ &dY_j^i = \mu(Y_j^i, t_j, \theta, b^i)dt + \Sigma(Y_j^i, \theta, b^i)dW^i(t), Y_0^i = y_0^i, i = 1, ..., M, \\ &b^i \sim i.i.d P_B(\cdot|\Psi) \end{aligned} \tag{3.29}$$

For an SDME model with measurement noise, the parameters to be estimated are (θ, Ψ, Ω). The likelihood function is then defined as:

$$L(\theta, \Psi, \Omega) = \prod_{i=1}^{M} \int_B (\int_Y P_{\underline{Z}}(\underline{z}^i | \underline{y}^i, \Omega) P_{\underline{Y}}(\underline{y}^i | b^i, \theta) d\underline{Y}^i) P_B(b^i | \Psi) db^i \tag{3.30}$$

where: $P_{\underline{Z}}(\underline{z}^i | \underline{y}^i, \Omega) = \prod_{i=1}^{n_i} P(z_j^i | Y_j^i, \Omega)$ is the conditional density of the observations z

given the diffusion Y. Here, for an SDME model with measurement noise, the difficulties encountered for the model without measurement noise with direct observations are still present and are exacerbated by integration with hidden trajectories $\underline{Y}^i$. The main objective is to estimate the parameters of the model, many estimation methods developed for an SDME model with measurement noise are suggested in the theory, in the following we will revise the most famous of them.

Firstly, we present the Bayesian approach adopted to an SDME model with a population approach proposed by: S Donnet 2010. Secondly, we present the extended Kalman Filter (EKF) method coupled with the First Order Conditional Estimate (FOCE) algorithm, this combined methodology was proposed by Overgaard et al. 2005 and Tornøe et al. 2005. Finally, we develop the Expectation-Maximization algorithm (E-M) in its stochastic version adopted to NLME model proposed by Walker 1996 and Kuhn and Lavielle 2005. As for the case of an SDME model without measurement noise, the choice of an estimation method depends on the availability of an explicit transition density.

Estimation methods:

Bayesian inference: Recurring difficulties of classical inference to maximize and have the likelihood function in a closed-form, that requires the existence of the transition density in explicit form, could be exceeded by using Bayesian inference. This method consists to set a prior distribution of the parameters and then to estimate the posterior distribution of these parameters (θ, Ψ, Ω) given the observations $\underline{z}$ according to the Bayes formula $P(\theta, \Psi, \Omega|\underline{Z}) \propto f(\underline{z}; \theta, \Psi, \Omega)\pi(\theta, \Psi, \Omega)$. So, the obtained estimators of (θ, Ψ, Ω) are characterized by statistical characteristics of the posterior distribution $p(\theta, \Psi, \Omega|z)$ (posterior mean or median). But, usually, the posterior distribution has no explicit expression, especially for nonlinear models, so iterative estimation procedures are needed. The recommended strategy is to use the Monte Carlo Markov Chain (MCMC) algorithm (See: Robert and Casella 2004) to generate a Markov chain sample from the posterior distribution as its marginal stationary distribution. This algorithm simply consists in generating alternately, at the iteration k:

(1) $b^{(k)}|\underline{Y}^{(k-1)}, \underline{z}, \theta^{(k-1)}, \Psi^{(k-1)}, \Omega^{(k-1)}$

(2) $\underline{Y}^{(k)}|b^{(k)}, \underline{z}, \theta^{(k-1)}, \Psi^{(k-1)}, \Omega^{(k-1)}$.

(3) $(\theta^{(k)}, \Psi^{(k)}, \Omega^{(k)})|\underline{Y}^{(k)}, b^{(k)}, \underline{z}$.

when it is assumed that there is no measurement noise and that the process is observed directly, the second step is removed. The Markov Chain produced by this algorithm has a stationary distribution $p(b, \underline{Y}, \theta^{(k)}, \Psi^{(k)}, \Omega^{(k)}|\underline{z})$, and after a high number of iterations, we get a sampled parameters $(\theta^{(k)}, \Psi^{(k)}, \Omega^{(k)})$ that are assumed to have this interest distribution $p(\theta, \Psi, \Omega|\underline{z})$ as its marginal posterior distribution, which responds to the goal. Moreover, as the conditional densities used in this algorithm are not explicit, we need to use the Metropolis-Hastings simulation algorithm.

Stochastic Expectation-Maximization (SAEM) algorithm: Stochastic Expectation-Maximization algorithm is another suggested method to estimate parameters of a nonlinear SDME model. This method was originally proposed by Dempster, Laird, and Rubin 1977, and treated in Wang 2007 for NLME models, then it was adopted to population SDEs by Donnet and Samson 2008, Donnet and Samson 2011 and Delattre and Lavielle 2013. This algorithm is a suitable method for an SDME model with noise measurement, it allows to avoid the problem of the integration in the likelihood function when it is possible to maximize the conditional expectation of the probability of the complete data $(\underline{z}, \underline{y}, b)$:

$$Q(\theta, \Psi, \Omega|\theta', \Psi', \Omega') = E[logP(\underline{z}, \underline{y}, b; \theta, \Psi, \Omega)|\underline{z}; \theta', \Psi', \Omega')] \tag{3.31}$$

The original Expectation-Maximization (EM) method is based on two steps: The first is the Expectation-step (E-step) which is an iterative procedure, where $Q_l(\theta, \Psi, \Omega)$ is evaluated at the l-iteration given the current value of the parameters $\theta_{l-1}, \Psi_{l-1}, \Omega_{l-1}$. Then, the second is the Maximization-step (M-step) where the likelihood function $Q(\theta, \Psi, \Omega|\theta'_{l-1}, \Psi'_{l-1}, \Omega'_{l-1})$ is maximized and $\theta_{l-1}, \Psi_{l-1}, \Omega_{l-1}$ are updated. Moreover, when the conditional distribution $P(\underline{y}, b|\underline{z}; \theta, \Psi, \Omega)$ is not explicit this algorithm cannot be applied, which is generally the case in population SDEs. So, the E-step is not an obvious step for the case of an SDME model, in this context, the stochastic version of the EM algorithm has been proposed by the authors above.

In the SAEM algorithm, the E-step is split into a simulation step (SM-step) and a stochastic approximation step (SA-step). The first simulates the non-observed data $(\underline{y}^{(l)}, b^{(l)})$ according to the conditional distribution $P(\underline{y}, b|\underline{z}; \widehat{\theta}_{l-1}, \widehat{\Psi}_{l-1}, \widehat{\Omega}_{l-1})$, this simulation step might not be easy to perform exactly, so other algorithms are required as: MCMC algorithm, when the transition density is explicit as in the Bayesian framework, or Particle MCMC (PMCMC) algorithm, which coupled the MCMC algorithm with particle filter techniques (See Andrieu, Doucet, and Holenstein 2010 and Donnet

and Samson 2011). These algorithms proved their convergence theoretically when the complete log-likelihood belongs to the exponential family with respect to the parameters (θ, Ψ, Ω). Then, the second step (SA-step) is based on the following approximated equation:

$$Q_l(\theta, \Psi, \Omega) = Q_{l-1}(\theta, \Psi, \Omega) + \eta_l[logP(\underline{z}, \underline{y}^{(l)}, b^{(l)}; \theta, \Psi, \Omega) - Q_{l-1}(\theta, \Psi, \Omega)] \quad (3.32)$$

where $(\eta_l)_{l\in\mathbb{N}}$ is a sequence of positive numbers decreasing towards zero. However, in the simulation step only the individual parameters b^i are simulated according to $P(b^i|z^i;\theta)$, the diffusion trajectories are not simulated but are directly used in this conditional distribution. Thus, this distribution cannot be directly simulated so the MCMC algorithm with a Metropolis-Hastings algorithm is recommended. But, to compute the probability of acceptance, the expression of $P(z^i|b^i, \Omega)$ is necessary, whereas, it does not have a closed-form. Thus, following all these constraints, an alternative method to avoid such problems has been proposed by Delattre and Lavielle 2013, based on the EKF algorithm described below to approximate the conditional likelihood to a Gaussian function.

EM-Bayesian inference: Here, we resume the Bayesian method seen above when the density of transition is not explicit. In this case, we consider an approximate model resulting from the Euler-Maruyama approximation of the SDE solution according to a size h step as proposed in Donnet and Samson 2011, then we perform the Bayesian inference on the model obtained as it was treated previously.

EM-SAEM: The SAEM algorithm can be adopted to the case of an SDME model with measurement noise but without an explicit transition density, using the approximate solution of the SDME model obtained by the Euler-Maruyama approximation. In this case, the transition density can be approximated by a Gaussian distribution if the time interval between two observations is small, if it is large we need to introduce a set of auxiliary latent data between every pair of observations in order to get a good approximation to the transition density as it was proposed by Donnet and Samson 2008. So, in practice, this introduction of unobserved data is generally recommended, which can slow down the convergence rate.

Extended Kalmen Filter (EKF) coupled with First-Order Conditional Estimation (FOCE) algorithm: This method is an approximate method to estimate a nonlinear SDME model with unknown transition density, it is constructed, as it is mentioned in the title,

by the combination of two techniques and that is valid only for models with measurement noise. The First-Order Conditional Estimation (FOCE) is a linearization tool that allows a nonlinear mixed-effect model to be compared to a linear model, see: Lindstrom and Bates 1990. Overgaard et al. 2005 and Tornøe et al. 2005, that proposed to adapt this method to the population SDEs model. For the Extended Kalman approach, it has been proposed by Tornøe et al. 2005 for a time-homogeneous SDME model with measurement noise, the main idea is to approximate the transition density by a Gaussian distribution, therefore, the combination of these methods could be effective for the inference problems for a nonlinear SDME model with an unknown transition density.

First, we apply the EKF method in order to approximate the individual likelihood functions by a Gaussian distribution, then, the FOCE algorithm is applied in order to facilitate and approximate the computation of the likelihood function of the population and parameter estimates. So, the likelihood function (3.30) of model (3.29) can be rewritten according to the idea of recursive conditioning in the following formula:

$$L(\theta,\Psi)=\prod_{i=1}^{M}\int_{B}[\prod_{j=1}^{N}P_Z(z_j^i,\Delta_j^i|z_{j-1}^i,b^i,\theta)]P_B(b^i|\Psi)db^i \tag{3.33}$$

In EKF algorithm, the transition density $P_Z(z_j^i,\Delta_j^i|z_{j-1}^i,b^i,\theta)$ is approximated by a Gaussian distribution, with mean $m^i_{j|1:j-1}(b^i,\Omega\theta)$ and variance $R^i_{j|1:j-1}(b^i,\Omega,\theta)$ which are not explicit when h is nonlinear, and depend on b^i,Ω and θ :

$$\begin{aligned} m^i_{j|1:j-1} &= E(z_j|z_{1:j-1}) = E(h(Y_j^i,\theta|z_{1:j-1}) \\ R^i_{j|1:j-1} = Var(z_j|z_{1:j-1}) &= Var(h(Y_j^i,\theta|z_{1:j-1})+g(Y_j^i,\Omega)^2 \end{aligned}$$

So, for the model (3.29), the approximate individual likelihood function can be rewritten as:

$$\begin{aligned} L_i^{EKF}(\theta,\Psi,\Omega;\underline{z}^i) &= \int_B[\prod_{j=1}^{n_i}\frac{exp(\frac{-1}{2}(z_j^i-m^{(i)}_{j|1:j-1})^T(R^i_{j|1:j-1})^{(-1)}(z_j^i-m^i_{j|1:j-1}))}{\sqrt{|2\pi R^i_{j|1:j-1}|}}] \\ P_B(b^i|\Psi)db^i &= \int_B e^{l_i(b^i)}db^i \end{aligned} \tag{3.34}$$

with l_i is the approximate conditional log-density of b^i given the observations.

$$\triangle l_i \approx -\sum_{j=1}^{n_i}(\bigtriangledown \varepsilon_{ij}^T R_{i(j|j-1)}^{-1} \bigtriangledown \varepsilon_{ij} - \Psi^{-1}) \tag{3.35}$$

Then, we approximate l_i by a second-order Taylor expansion (using Laplacian approximation), the gradient of l_i with respect to the random effects will vanishes because in FOCE algorithm the expansion is evaluated around the true minimum $\widehat{b^i}$ of l_i. So, we get the following approximated likelihood from EKF algorithm and Taylor expansion:

$$logL^{EKF,FOCE}(\theta, \Psi, \Omega; \underline{z}) = \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}} | - H(l_i)|^{\frac{-1}{2}} exp(l_i)_{|_{\widehat{b^i}}} \tag{3.36}$$

So, it remains only the computation of the Hessian term $H(l_i)$ to deduce the likelihood function of the population, which cannot be calculated exactly but it could be approximated using the FOCE algorithm. For more information on the implementation of the FOCE method in computer software (See Tornøe et al. 2005, Mortensen et al. 2007 and Klim et al. 2009). This method is not yet defended theoretically since no theoretical convergence has been proved.

3.7 Conclusion

In this chapter, we have proposed an estimate procedure of a mixed effects model containing stochastic differential equations, known by the SDME models, to get the likelihood function in an approximate closed form to obtain the ML estimators. Then, we have presented a review on estimation methods existing in the literature for more knowledge on the estimation procedures already existing in this context. The proposed method will be evaluated, in Chapter 5, using simulation studies on two examples of SDME models: the two-dimensional Ornstein-Uhlenbeck process and stochastic minimal model describing the glucose-insulin kinetics.

Generally, in models with SDEs instead of ODEs with random effects, the estimation of parameters still not obvious even for one individual (one trajectory), because of the

difficulties in deriving the transition densities. These difficulties become more interesting when using the population approach which treats the entire population simultaneously. In fact, the derivation of the exact density is not always possible for a stochastic and continuous process in an SDME model, so looking for an approximation is an important step and requires expensive calculation. This task is very interesting to give good results with good statistical properties of the estimators obtained by maximizing the likelihood function. Here, we have proposed an approximation method to obtain the transition density in a closed form, based on the Risken approximation for the formal solution of the Fokker-Planck equation proposed by Risken in Risken 1996.

In the theory, many methods have been proposed depending on the observed process in the model whether it includes measurement noise or not, we can summarize the whole described procedure, as mentioned above, in the following way: (1) For an SDME model without measurement noise: i) if the exact transition density exists, this is the simplest but the rarest case, model parameters could be estimated by giving the exact ML estimators; ii) otherwise, we propose to use our estimation methodology proposed in this section, or the Hermit expansion, or the solution of the Kolmogorov equations as approximate methods to estimate the transition density. (2) When the model includes measurements noise: i) Bayesian inference and Stochastic Expectation-Maximization (SAEM) algorithm are recommended when the transition density is known ii) if not, these methods are applied on the approximated solution derived by Euler-Maruyama (EM) scheme or replaced by Extended Kalmen Filter (EKF) coupled with First-Order Conditional Estimation (FOCE) algorithm. Indeed, the choice and use of each of these methods is restricted according to the case of the application of each of them. For example: when the model includes measurements noise, the Hermite expansion cannot be directly used while the EKF method coupled to FOCE is suitable, however, the Bayesian approach is adapted for any model (with or without measurement noise) and when also the transition density is not known. In addition, for the EM approximation, it must be performed with the introduction of latent data between each pair of observations to increase the convergence of the algorithm and the quality of the estimators.

Moreover, the statistical inference for SDME models implies the problem of the numerical evaluation of the integral given random effects in the likelihood function, which becomes complicated especially when the model contains more than two random parameters. In the literature, several methods have been proposed and tested for the approximation of the integral (See references in the introduction).

Finally, we notice that researches are still in progress for more efficient and reliable

estimation methods for SDME models, as these models are still of interest to scientists who believe that, this type of model will have increasing popularity, as it combines the right characteristics of mixed effects theory, with the possibility of considering the system noise in the intra-subject dynamics, thus leading to a flexible and powerful modelling approach.

Chapter 4

Insulin sensitivity modelling

Contents

4.1 Background

4.1.1 About glucose and insulin:

Insulin is a hormone naturally secreted by the pancreas. When food is ingested, carbohydrates are broken down into glucose. Glucose, in turn, serves as a source of energy for the body. The pancreas produces and releases insulin to help the body use and

/ or store this glucose. Insulin works in combination with other hormones including amylin and glucagon. There are two main types of insulin: Slow-acting insulin: which balances blood sugar levels throughout the day. It is the insulin that the body needs to function properly, and Fast-acting insulin (or rapid insulin) is the insulin the body needs to cover carbohydrate intake from meals. This insulin lowers blood sugar in hyperglycemia (high blood sugar level).

4.1.2 Diabetes

Diabetes mellitus is a common metabolic disease that spreads in an interesting way among people around the world. The study of this disease is based on observations of the body's responses to the insulin produced by the pancreas, that should be used properly to take advantage of glucose in the tissues. The diabetes disease is one of the most prevalent diseases between individuals for both sexes, its affectation degree depends on certain characteristics which depend on each individual and manifests in different ways between individuals according to two types of diabetes (See: Alberti and Zimmet 1998): Type 1 Diabetes (T1D): when the body fails to produce insulin at all, which involves the injection of insulin to regulate the level of glucose in the blood, and Type 2 Diabetes (T2D): when the cells do not react to the secreted insulin, because they do not use insulin properly or if the produced insulin is insufficient.

The glucose is produced mainly by the liver, distributed and used both in the blood system and red blood cells on the one hand; constituting insulin-independent; and in muscles and tissues on the other hand; constituting insulin-dependent. Insulin is secreted by the pancreatic beta cells and then it enters the circulation of the system after liver degradation, then it is eliminated primarily by the kidneys, these controlling interactions are called insulin sensitivity and beta cell sensitivity. However, the incompatibility or insufficiency of the produced insulin causes a dysfunction of the insulin role in favour of glucose, resulting in an inability to remove glucose from the blood at the normal rate. These abnormalities can be explained by a low sensitivity to insulin (See: Pacini and Bergman 1986), which means the insulin's ability to reduce glucose levels in the muscles, liver and tissues, or by a weak glucose effectiveness that means the ability of glucose to improve its own elimination at the basal insulin level, or by the failure of the pancreatic β-cells to secrete insulin in response to glucose stimuli. The quantitative assessment of these critical factors is possible from the minimal model (See: Bergman et al. 1979a) that could improve the identification and diagnosis of the glucose-insulin kinetics in blood to describe the diabetes disease and its corresponding

treatments (See: Martin et al. 1992), the aim being therefore to investigate how good estimates of these factors can be obtained.

The idea is to use the minimal model and to explore the Intra Venous Glucose Tolerance Test (IVGTT) dataset, which includes measurements of glucose and insulin concentrations in the blood before and after an intravenous glucose injection at different time points, in order to model and estimate the key diabetes parameters. The minimal model is composed by ordinary differential equations (ODEs) sets describing the glucose-insulin kinetics simultaneously, it is based on a compartment model structure similar to what is commonly used for modelling within Pharmacokinetics (PK). In Pacini and Bergman 1986, the glucose and insulin kinetics were described separately by two sets of differential equations, but this approach leads to estimate, separately, the parameters which often provide unrealistic parametric estimates, and in Pillonetto et al. 2002, the Bayesian approach is adopted to estimate glucose kinetics parameters by considering the insulin sensitivity parameter as known, however, this assumption leads to the loss of important information. So, it is of interest to couple the dynamics of the kinetics of glucose and insulin concentrations in a single model as assumed in the minimal model structure. Nevertheless, as pointed out by the authors in De Gaetano and Arino 2000, this coupled model can lead to a very poorly posed and complicated parameter estimation problem, this is the targeted purpose of the present work. In the following, the estimated approach proposed above will be applied to the minimal model in its stochastic version, where the empirical results will be reported later in Chapter 5.

4.1.3 Insulin sensitivity

Insulin sensitivity is an important parameter for diabetes diagnosis. It is one of the parameters that describe the insulin-glucose kinetics in the body and predict whether a person has a high risk of developing Type 2 diabetes. Moreover, it indicates how the body responds to the secreted insulin by measuring the ability of insulin to increase the efficiency of glucose to cells or tissues such as muscles and liver. This parameter therefore makes it possible to describe the body's sensitivity to the effects of insulin. For a person with a normal insulin sensitivity, a smaller amount of insulin is sufficient to lower his blood glucose level than someone who has low sensitivity, whereas, a person with low sensitivity requires larger amount of insulin either secreted by the pancreas or from injections to keep blood glucose stable in a normal rate. Moreover, when the insulin sensitivity is lower we can talk about insulin resistance. So, this

parameter is very important to determine the abnormalities in the insulin metabolism, however, it can be quantified only if we are able to assess insulin action in the pancreas and to measure the glucose and insulin entering and leaving the system. In this work, we mainly aim to quantify this parameter by estimating minimal model parameters describing the glucose-insulin control system (See next paragraphs).

4.1.4 C-peptide and insulin kinetics

The C-peptide is a substance made in the pancreas and secreted by beta cells, passes through the liver before appearing in plasma without being extracted by the liver. One of the major problems encountered in estimating insulin sensitivity is the measurement of insulin secretion, because it is not possible to infer pancreatic secretion from data on plasma insulin concentration; it is only possible to derive its component appearing in plasma or the secretion of posthpatic insulin, which is approximately equal to 50% of the pancreatic secretion. However, the C-peptide pancreatic secretion coincides with insulin secretion because it is secreted equimolarly to insulin. In other words, plasma C-peptide concentration offers o classical way to estimate the insulin secretion during perturbation from plasma concentration measurements. So, the noted constraint can be easily overcome if we have an available data on C-peptide measurements.

4.2 Minimal model of glucose-insulin kinetics

4.2.1 Presentation

The minimal model is based on a compartmental system outcome from individual compartment analysis, which is one of the methods used in PK analysis as mentioned before (See General Introduction). It was proposed in the late 1970s by Bergman et al. 1979b and later developed in Toffolo, De Grandi, and Cobelli 1995 and Bergman, Phillips, and Cobelli 1981, it describes the glucose-insulin kinetics and the dynamic of these processes, in order to illustrate the mechanisms of diabetes disease, where the measurements are based on data from IVGTT. Moreover, the minimal model was originally specified for a single individual and does not combine several individuals, however, using population-based approach with the aim of estimating the metabolic portrait for a whole population can be very useful in the study of diabetes, with the advantage of estimating the metabolic portrait for a whole population.

As mentioned earlier, diabetes is one of the most prevalent diseases in individuals, the degree and type of its affectation varies from individual to individual and depends on certain individual characteristics, which implies that the concepts of stochastic modelling with random effects could be a good approach for diabetes modelling. Moreover, as seen before, there are two principal types of diabetes; T1D due to the insufficient insulin production or to the fact that the cells do not respond to the secreted insulin, and T2D where the patients tend to have substantially lower insulin sensitivity than healthy individuals. Thus, to model the T2D, we observe how a person's body responds to insulin in the process of transporting glucose to tissues, by measuring his insulin sensitivity. In this chapter, we deal with the estimation of the minimal model which represents a powerful model describing the glucose-insulin kinetics in three differential equations simultaneously, see the mathematical formulation of the model in: Bergman et al. 1979b; Bergman, Phillips, and Cobelli 1981; Toffolo, De Grandi, and Cobelli 1995; Cobelli et al. 2009. Therefore, it might already be clear that the model should take into account the variability between and within individuals by containing both fixed and random effects, since the diabetes disease analysis takes into account the response of each individual according to his own parameters as well as others which are common and which describe the process of glucose-insulin for the entire population, see Figure (4.1).

At first, glucose and insulin concentrations in the blood are described by two sets of differential equations (See: Pacini and Bergman 1986); at a rate p_1, glucose leaves and enters the glucose space in proportion to the difference between the plasma glucose concentration $G(t)$ and the basal plasma concentration G_b, which is known and represent the pre-injection glucose level for each individual. Therefore, the parameter p_1 represents the glucose's own ability to be eliminated in muscles, liver, and tissues independently of insulin and which is called glucose efficiency and denoted by S_G. Then, the glucose disappears from the glucose space at a rate proportional to insulin concentration in the insulin compartment $X(t)$, which represents the dynamic of insulin response according to the two rates p_2 and p_3. These two parameters represent, respectively, the decreased glucose absorption capacity in tissues and its increased insulin dependency, the insulin sensitivity is defined by combining these two rates where $S_I = p_3/p_2$, representing insulin's ability to increase the net glucose utilization, Bergman et al. 1979a. For insulin secretion $I(t)$, it is secreted by the pancreas independently of the glucose concentration, and proportional to a rate n to its own level already in the body and to the glucose level deferred from a threshold h at a rate γ when $G(t)$ is above h, the insulin secretion then depends not only on the hyperglycemia level but also to the time spent since glucose injection.

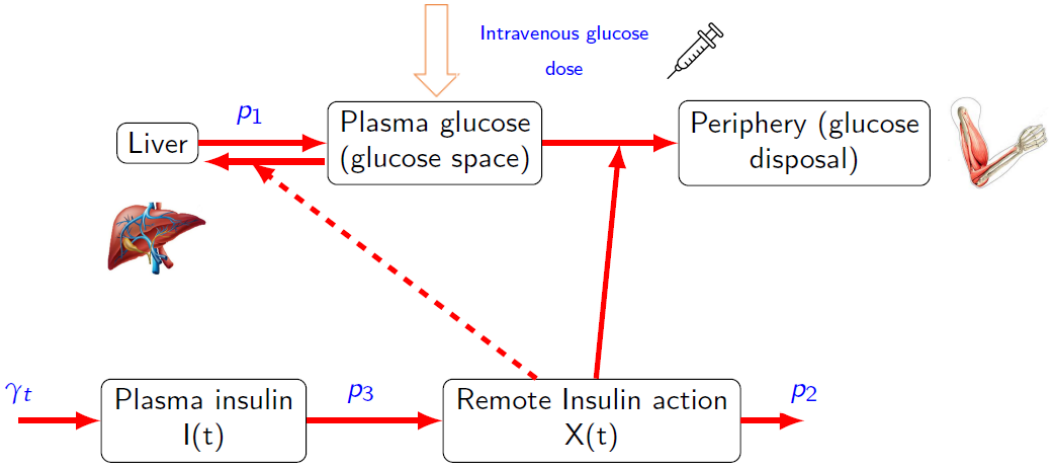

FIGURE 4.1: The minimal model: Scheme of the glucose-insulin system.

Beforehand, in the rest of our work, the population is considered to be non-diabetic, so that the reactions of the subjects do not influence the results in order to validate the estimation procedure proposed while avoiding any confusion. However, nothing prevents studying a diabetic population in a real application.

4.2.2 Stochastic minimal model

As mentioned earlier, the insulin and glucose processes, like all actual pharmacological processes, could not be deterministic because they are exposed to many other unknown factors. So, stochastic modelling with random individual parameters seems to be a good analytical tool for diabetes disease modelling and insulin sensitivity estimation for each person. In this study, we use measurements obtained from the intravenous glucose tolerance test (IVGTT), see: De Gaetano and Arino 2000 for a mathematical modeling of the test where glucose and insulin concentrations in plasma are subsequently sampled after an intravenous glucose injection), where a small amount of glucose is administered intravenously to the subjects and then the concentration of glucose and insulin in plasma are observed after different points of time, in this context, we refer the reader to De Gaetano and Arino 2000 which describes the mathematical modeling of the IVGTT.

In the literature, including the above-mentioned papers, the minimal model has been considered as a deterministic model with only measurement errors, but we believe that the dynamic of glucose and insulin processes modelized by the ODEs may not correspond to the real dynamics of these processes in the body. So, there is an increasing need to extend the deterministic minimal model to its stochastic version using the

stochastic differential equations by adding stochastic components of Brownian motion to the system of the ODEs. In this chapter, we deal with the use of a stochastic differential mixed effects model for the diabetes diseases modelling, using the minimal model in its stochastic version.

The mathematical formulation of the system is presented by three differential equations describing the glucose-insulin kinetics according to the parameters appearing in Figure (4.1), and describes through three SDEs how a person's body responds to insulin in the process of transporting glucose to tissues, and defines the mechanism of the glucose-insulin kinetics (See: Bergman et al. 1979b; Bergman, Phillips, and Cobelli 1981; Toffolo, De Grandi, and Cobelli 1995; Cobelli et al. 2009).

From the mentioned literature and Figure (4.1), the glucose-insulin disposal can be presented, with respect to time, by the following nonlinear stochastic differential equations, perturbed by the stochastic components $\sigma_1 dw_1(t)$, $\sigma_2 dw_2(t)$ and $\sigma_3 dw_3(t)$:

$$\begin{aligned} dG(t) &= [-(p_1 + X(t))G(t) + p_1 G_b]dt + \sigma_1 dw_1(t),\ G(0) = G_0 \\ dX(t) &= [-p_2 X(t) + p_3(I(t) - I_b)]dt + \sigma_2 dw_2(t),\ X(0) = 0 \end{aligned}$$

where $G(t)$ and $I(t)$ are, respectively, the concentration of glucose and insulin at time t in the blood. G_b and I_b indicate the basal level of glucose and insulin concentration before the glucose injection. This injection will cause a disturbance of the concentrations according to the mechanism described in these equations, these values are assumed known for each individual. Also, G_0 and I_0 are the theoretical measure of the concentrations at glucose injection moment at the beginning of the experiment.

For the insulin secretion, measured by $I(t)$, it is secreted by the pancreas proportionally to its own level already in the body at a rate n, and also to the glucose level deferred from a threshold h at a rate γ, and to the time spent since glucose injection. Thus, the kinetic of the insulin is directly dependent on time and can be presented by the following SDE:

$$dI(t) = [-n(I(t) - I_b) + \gamma(G(t) - h)t]dt + \sigma_3 dw_3(t), I(0) = I_0$$

where I_0 is the theoretical initial insulin concentration in plasma.

Finally, the stochastic minimal model based on the population approach in the sense of Itô formula; and reparametrized by S_G and S_I; can be defined as:

$$dY^i(t) = \xi(Y_t^i, t, \theta, b^i)dt \; + \; \Sigma dW(t) \; ; \; Y_0^i = \begin{pmatrix} G_0^i \\ 0 \\ I_0^i \end{pmatrix} \tag{4.1}$$

where : $Y^i(t) = \begin{pmatrix} G(t)^i \\ X(t)^i \\ I(t)^i \end{pmatrix}$, and $\xi(Y_t^i, t, \theta, b^i) = \begin{pmatrix} -(S_G^i + X(t)^i)G(t)^i + S_G^i G_b^i \\ -p_2(X(t)^i + S_I^i(I(t)^i - I_b^i)) \\ -n(I(t)^i - I_b^i) + \gamma(G(t)^i - h)t \end{pmatrix}$ which represents the drift term , and the diffusion term Σ is a diagonal matrix of elements σ_1, σ_2 and σ_3.

The ξ and Σ fulfil the Lipschitz (2.5) and linear growth (2.4) conditions (See Appendix A.1 for more details). The parameters $S_G^i, p_2, S_I^i, n, \gamma, h, G_0^i$ and I_0^i are unknown in the model and should be estimated. The parameters $S_G^i, S_I^i,\ I_0^i$ and G_0^i are assumed random, because they represent individual parameters that change from an individual to another. Indeed, each subject has its own insulin sensitivity S_I^i which allows to know if the cells of his body react correctly or not to the insulin; and if the insulin produced by the pancreas is sufficient or not, that can make some people with T2D and others without diabetes. Also, for glucose effectiveness S_G^i, which represents the glucose's own ability to be eliminated independently of insulin, it is unique to each individual and changes from a person to another, as well as for the measurement of glucose and insulin concentrations. For the rest of the parameters, we consider them fixed since they describe the common side of the glucose-insulin kinetics for the entire population. So, we have the following random effects vector $b^i = (S_G^i,\ S_I^i,\ I_0^i,\ G_0^i)$, and we assume that :

$$S_G^i \sim \mathcal{N}(\mu_{S_G}, \sigma_{S_G}),\ S_I^i \sim \mathcal{N}(\mu_{S_I}, \sigma_{S_I}),\ I_0^i \sim \mathcal{N}(\mu_{I_0}, \sigma_{I_0}),\ G_0^i \sim \mathcal{N}(\mu_{G_0}, \sigma_{G_0}). \tag{4.2}$$

Random effects are assumed to be independent and identically distributed with a multinormal joint density function, with the mean $\vartheta = (\mu_{S_G}, \mu_{S_I}, \mu_{I_0}, \mu_{G_0})$, and the covariance matrix $\phi = diag(\sigma_{S_G}, \sigma_{S_I}, \sigma_{I_0}, \sigma_{G_0})$, so we have: $\Psi = (\mu_{S_G}, \mu_{S_I}, \mu_{I_0}, \mu_{G_0}, \sigma_{S_G}, \sigma_{S_I}, \sigma_{I_0}, \sigma_{G_0})$ and $\theta = (p2, n, \gamma, h, \sigma_1, \sigma_2, \sigma_3)$.

Traditionally, the minimal model has been analysed in a deterministic setup with only error terms on the measurements, however, the statistical inference of its stochastic version is not an obvious procedure, since the transition density of the stochastic process is usually unknown.

In this work, the processes are considered stochastic instead of deterministic because, as explained before, the actual pharmacological processes are always exposed to influences that are not completely understood or that it is impossible to model explicitly, and ignoring these phenomena in the modeling may affect the estimation result. Moreover, the minimal model was already described for a single individual and does not combine several individuals, but we are interested in estimating the metabolic portrait of an entire population instead of a single individual using the population approach. For this purpose, two types of parameters have been incorporated in the model: fixed effects to capture general and common behavior for the whole population and random effects varying between individuals to account for individual deviation. However, the obtaind mixed-effects model, considering both glucose and insulin dynamics simultaneously, is an extremely poorly estimated estimation problem where the reconstruction was most often performed using non-linear least squares techniques separately for each entity.

4.2.3 Insulin sensitivity estimation

From the description of the minimal model above, we deal then here with a time-inhomogeneous NLME model with SDEs describing the glucose-insulin kinetics; see: Egorov, Li, and Xu 2003 for the implementation of SDE time-inhomogeneous model, and Picchini, Ditlevsen, and De Gaetano 2008 where the maximum likelihood estimation for a time-inhomogeneous stochastic differential model of glucose dynamics was treated. We notice that the measurements in the model (4.1) are assumed observed directly without measurement errors.

So, we wish to estimate (θ, Ψ) given the observations $\underline{y} = (\underline{y^1}, ..., \underline{y^M})$ from model (4.1). By using the approximated transition density (3.20) we get the following approximated likelihood function for model (4.1):

$$\begin{aligned}\mathcal{L}^{(a)}(\theta,\Psi) \;=\; & \prod_{i=1}^{M}(2\pi)^{-2}(\sigma_{SG}\sigma_{SI}\sigma_{I0}\sigma_{G0})^{\frac{-1}{2}}\prod_{j=1}^{n_i}((2\sqrt{\Pi\Delta_j})^{-3})[\sigma_1\sigma_2\sigma_3]^{-\frac{1}{2}})\int_{R^4} exp(\sum_{j=1}^{n_i}[\frac{-1}{4\Delta_j} \\ & [\frac{1}{\sigma_1}(A^i_{1j})^2+\frac{1}{\sigma_2}(A^i_{2j})^2+\frac{1}{\sigma_3}(A^i_{3j})^2)]-\frac{1}{2}(\sigma^{-1}_{S_G}(S^i_G-\mu_{S_G})^2-\sigma^{-1}_{S_I}(S^i_I-\mu_{S_I})^2-\end{aligned}$$

$$\sigma_{I_0}^{-1}(I_0^i-\mu_{I_0})^2-\sigma_{G_0}^{-1}(G_0^i-\mu_{G_0})^2))dS_G{}^i dS_I{}^i dI_0{}^i dG_0{}^i \tag{4.3}$$

with:

$$A_{lj}^i=[(Y_j^i)_l-(Y_{j-1}^i)_l-\xi_l(Y_{j-1}^i,t_{j-1},\theta,b^i)\Delta_j^i],\ l=1,2,3.$$

We have no closed form solution to this integral, so exact estimators of θ and Ψ are unavailable. Therefore, we use the Laplace approximation method described in Chapter 3 to obtain an approximate closed form of the log-likelihood function $log(L^{(a)}(\theta,\Psi))$ for the model (4.1); then by applying the GA, we get the approximate estimators $\hat{\theta}$ and $\hat{\Psi}$.

Otherwise, using (3.21) instead of (3.20), we obtain a likelihood function with an integral that we can solve analytically, and which allows to obtain the exact ML estimators of the parameters (θ,Ψ) giving a system of 15 gradient equations (See Appendix A). Indeed, to have several forms of solution for the motion equation (3.16) of the transition density, allows to choose according to each example the best one which leads to obtain the exact estimators. The formulas of the fifteen exact estimators of (θ,Ψ) are obtained analytically by deriving the likelihood gradient. We give below some of the ML estimators representing the main parameters to estimate the insulin sensitivity, other details and estimators are reported in Appendix A.1.2:

$$\widehat{\mu}_{S_I}=\frac{\sum_{i=1}^M \frac{\widehat{E}_{21}}{2\widehat{E}_3}}{M-\sum_{i=1}^M \frac{1}{2\widehat{E}_3\widehat{\sigma_{S_I}}}} \qquad \widehat{\sigma}_{S_I}=\sum_{i=1}^M\left(\frac{-4\widehat{\mu}_{S_I}\widehat{E}_2\widehat{E}_3+\widehat{E}_2^2}{4M\widehat{E}_3^2}+\frac{1}{2M\widehat{E}_3}\right)+\widehat{\mu}_{S_I}^2,$$

$$\widehat{\mu}_{S_G}=\sum_{i=1}^M\frac{\widehat{d}_1}{2M\widehat{a}_1} \qquad \widehat{\sigma}_{S_G}=\widehat{\mu}_{S_G}^2+\frac{1}{2M\widehat{a}_1}+\sum_{i=1}^M\frac{\widehat{d}_1^2-4\widehat{d}_1\widehat{a}_1\widehat{\mu}_{S_G}}{4M\widehat{a}_1^2}$$

where:

$$\widehat{E}_{21}=\sum_{j=1}^{n_i}\frac{-1}{2\Delta_j\widehat{\sigma}_{22}}(X_j-X_{j-1}+\widehat{p}_2X_j\Delta_j)\widehat{p}_2(I_j-I_b)\Delta_j;\ \widehat{E}_3=\sum_{j=1}^{n_i}\frac{1}{4\Delta_j\widehat{\sigma}_{22}}(\widehat{p}_2(I_j-I_b)\Delta_j)^2+\frac{1}{2\widehat{\sigma}_{SI}};\ \widehat{E}_2=\sum_{j=1}^{n_i}\frac{-1}{2\Delta_j\widehat{\sigma}_{22}}(X_j-X_{j-1}+\widehat{p}_2X_j\Delta_j)\widehat{p}_2(I_j-I_b)\Delta_j+\frac{\widehat{\mu}_{S_I}}{\widehat{\sigma}_{S_I}};\ \widehat{d}_1=\sum_{j=2}^{n_i}\frac{-C_5C_6}{2\Delta_j\widehat{\sigma}_{11}}-$$

$\frac{C_3}{4\Delta_1\widehat{\sigma}_{11}} - \frac{C_4}{8\Delta_1\widehat{\sigma}_{11}\widehat{a}}$; $\widehat{a}_1 = \sum_{j=2}^{n_i} \frac{C_6^2}{4\Delta_j\widehat{\sigma}_{11}} + \frac{1}{2\widehat{\sigma}_{S_G}} + \frac{C_2}{(2\sigma_{G_0} + 4\Delta_1\sigma_1)}$; $C_2 = (G_1 - G_b)^2\Delta_1^2$; $C_3 = 2(G_1 - G_b)\Delta_1^2 X_1 G_1 + 2G_1(G_1 - G_b)_1$; $C_4 = -2(G_1 - G_b)\Delta_1$; $C_5 = G_j - G_{j-1} + G_j X_j \Delta_j$; $C_6 = (G_j - G_b)\Delta_j$; $a = \frac{1}{4\Delta_1\sigma_{11}} + \frac{1}{2\sigma_{G_0}}$.

Chapter 5

Implementation issues and numerical applications

[1]

Contents

In this section, we deal with numerical simulations of SDME models through two examples: In the first one, as a famous stochastic process, we consider trajectories of the Ornstein–Uhlenbeck process incorporating random effects, and the second one is the stochastic version of the minimal model that we have already presented in previous sections. The main of this section is to check the feasibility and effectiveness of the estimation procedure proposed above, and to apply the theoretical approaches and tools that we have presented and defined previously, when it is necessary. Data points from the trajectories were retrieved and the parameters were estimated on the obtained artificial data sets, which are generated with different sample sizes, e.g. (M=10, 20, 40

[1] The main results in this chapter are published in Fadwa, El Maroufy, and Mousse 2020

subjects and m=10, 20, 60 observations collected on each subject). Then, the applications with real data are given later in this chapter.

For each SDME model, artificial data sets of dimensions $m \times M$ were generated using different sets of parameters and different values of M and m, and the corresponding exact and/or approximated ML estimators were obtained. Finally, the purpose in this chapter is to stress the usefulness of the approximate closed-form of the transition density proposed above, either when the ML estimators are exact or approximated. Moreover, we notice that the numerical simulations of thousands of trajectories of the process are not required in each step of the optimized program, however, the search for optimization points can be considered as moderately expensive computationally since it takes a few hours (18 hours) depending on the size (M, m) of the sample. In our instances the parameter estimates were all obtained using a Matlab program.

5.1 Example 1: The two-dimensional Ornstein–Uhlenbeck process

To apply the proposed methodology and evaluate its effectiveness, we consider the two-dimensional stochastic OU process that is very useful in pharmacokinetic/pharmacodynamic studies, biology (See: Favetto and Samson 2010), physics, engineering, finance, and neuroscience applications (Picchini, Ditlevsen, and De Gaetano 2008; Ditlevsen and De Gaetano 2005a). Indeed, the choice of this process is due to the fact that it is one of the few known multivariate SDME models with known transition density. For this reason, we choose the OU process to evaluate the methodology presented above, by performing a comparison study between the results obtained using the proposed transition density in (3.20) and those obtained using its exact density. The model is defined as follows:

$$\begin{aligned} dY^{(1)i}(t) &= -(\beta_{11}b^i_{11}(Y^{(1)i}(t)-\alpha_1)+\beta_{12}b^i_{12}(Y^{(2)i}(t)-\alpha_2))dt+\Sigma_{11}dW^{(1)i}(t),\ Y^i_0=y^{(1)i}_0 \\ dY^{(2)i}(t) &= -(\beta_{21}b^i_{21}(Y^{(1)i}(t)-\alpha_1)+\beta_{22}b^i_{22}(Y^{(2)i}(t)-\alpha_2))dt+\Sigma_{22}dW^{(2)i}(t),\ Y^i_0=y^{(2)i}_0 \end{aligned} \tag{5.1}$$

$i = 1, ..., M$

With: $Y^i(t) = \begin{pmatrix} Y^{(1)i}(t) \\ Y^{(2)i}(t) \end{pmatrix}$; $\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$; $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$; $\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$; $W^i(t) = \begin{pmatrix} W^{(1)i}(t) \\ W^{(2)i}(t) \end{pmatrix}$; $Y^i(0) = \begin{pmatrix} Y^{(1)i}(0) \\ Y^{(1)i}(0) \end{pmatrix}$ and $b^i = \begin{pmatrix} b^i_{11} & b^i_{12} \\ b^i_{21} & b^i_{22} \end{pmatrix}$ where: $b^i_{ll'}$ $i.i.d$ $\sim \Gamma(r_{ll'}, r^{-1}_{ll'})$, $l, l' = 1, 2$; $i = 1, ..., M$.

We rewrite the system in matrix notation under the Itô formula; we denote by $(\cdot)$ the elementwise multiplication:

$$dY^i(t) = \beta \cdot b^i(\alpha - Y^i(t))dt + \Sigma dW^i(t),\ Y^i_0 = y^i_0,\ i = 1, ..., M. \tag{5.2}$$

Here, the random effects b_i are a matrix and not a vector in order to have a uniform dimension in writing of the (5.1) and are assumed mutually independent and independent of Y^i_0 and W^i. The fixed parameters vector is $\theta = (\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \alpha_1, \alpha_2, \Sigma_{11}, \Sigma_{22})$ and the population parameters vector is $\Psi = (r_{11}, r_{12}, r_{21}, r_{22})$. The exact transition density of model (5.2) for a given realization of the random effects is a bivariate Normal:

$$P_Y(Y^i_{t_j}, \Delta^i_j | Y^i_{t_{j-1}}, b^i, \theta) = (2\pi)^{-1} |\varsigma|^{\frac{-1}{2}} exp\left(\frac{-(Y^i_{t_j} - \mu)' \varsigma^{-1} (Y^i_{t_j} - \mu)}{2}\right), \tag{5.3}$$

with mean vector $\mu = \alpha + (Y^i_{t_{j-1}} - \alpha) exp(-(\beta.b^i)\Delta^i_j)$ and covariance matrix $\varsigma = \tau - (exp(-(\beta.b^i)\Delta^i_j)\ \tau\ exp(-(\beta.b^i)'\Delta^i_j)))$, where:

$$\tau = \frac{1}{2tr(\beta.b^i)|\beta.b^i|}(|\beta.b^i|\Sigma\Sigma' + (\beta.b^i - tr(\beta.b^i).I)\Sigma\Sigma'(\beta.b^i - tr(\beta.b^i).I)'),$$

We assume that the matrices $\beta.b^i$ and Σ have full rank and the real parts of the eigenvalues of $\beta.b^i$ are positive definite in order that a stationary solution to (5.2) exists. Under these assumptions, we derive from (5.1) and (3.20) the following approximated transition density of Y:

$$P^{(a)}_Y(Y^i_{t_j}, \Delta^i_j | Y^i_{t_{j-1}}, b^i, \theta) = (2\sqrt{\Pi\Delta_j})^{-2}(\Sigma_{11}\Sigma_{22})^{-\frac{1}{2}} * exp(-\frac{1}{4\Delta_j}(\Sigma^{-1}_{11})[Y^{(1)i}_{t_j} - Y^{(1)i}_{t_{j-1}} +$$

$$(\beta_{11}b^i_{11}(Y^{(1)i}_{t_{j-1}} - \alpha_1) + \beta_{12}b^i_{12}(Y^{(2)i}_{t_{j-1}} - \alpha_2))\Delta^i_j]^2 + (\Sigma^{-1}_{22})[Y^{(2)i}_{t_j} - Y^{(2)i}_{t_{j-1}} + (\beta_{21}b^i_{21}(Y^{(1)i}_{t_{j-1}} - \alpha_1) + \beta_{22}b^i_{22}(Y^{(2)i}_{t_{j-1}} - \alpha_2))\Delta^i_j]^2). \tag{5.4}$$

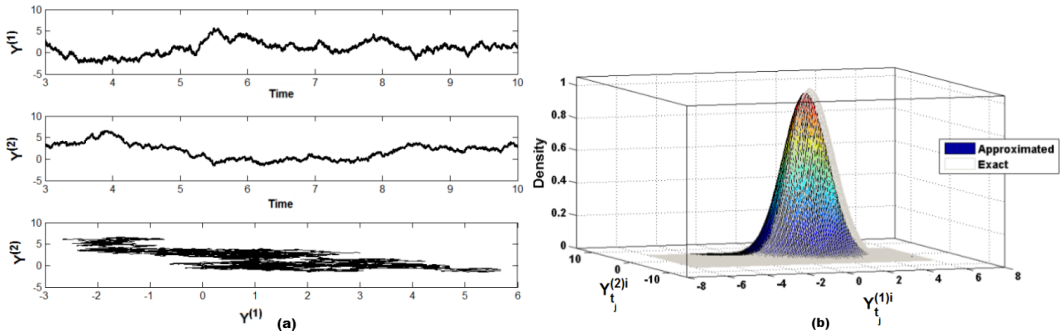

FIGURE 5.1: A sample path of the OU process is in the third graph of (a) for the given parameters set with the initial condition: $Y_0 = (3, 3)$ and time interval $[3, 10]$; and the transition density for a transition from Y_j to Y_{j+1} is in (b).

In Figure (5.1), the simulation of the OU process is presented using the Euler scheme (See: Klöden and Platen 1992) and the following set of parameters: $(\beta_{11} = 2.8, \beta_{12} = 2.5, \beta_{21} = 1.8, \beta_{22} = 2, \alpha_1 = 0.8, \alpha_2 = 1.5, \Sigma_{11} = 0.3, \Sigma_{22} = 0.5, r_{11} = 45, r_{12} = 100, r_{21} = 100, r_{22} = 125)$ with a time step of $\Delta_j = 0.001$; then we represent in (b) the graph of the two transition densities given by (5.3) and (5.4) for Y_j to Y_{j+1} using the same set of parameters and time step.

5.1.1 Simulation study

In this simulation study, we generate 1000 artificial datasets of dimension $2 \times (m + 1) \times M$ from (5.2), where M is the number of subjects and m presents the number of repetitions of the experiment for each subject; then we estimate the parameters using the proposed approximated method and we obtain 1000 sets of parameter estimates. The observations are obtained by linear interpolation from simulated trajectories using the Euler-Maruyama scheme with step size equal to 10^{-3}.

By plugging (5.3) in (3.4) and (5.4) in (3.6), we obtain a huge expression of the likelihood function but without a closed solution of integrals, so the exact estimators of θ and Ψ are unavailable. Therefore, in both cases, either using the exact or the approximated transition density, the Hessian matrix in Laplace approximation can be obtained analytically after a tedious calculation; then we apply the GA to obtain parameter estimates, which was expansively time-consuming to get the results because of the long and complex expressions formulas.

True values				**Mean and (Std) ($M = 40, m = 10$)**			
β_{11}	β_{12}	β_{21}	β_{22}	$\tilde{\beta}_{11}$	$\tilde{\beta}_{12}$	$\tilde{\beta}_{21}$	$\tilde{\beta}_{22}$
2.8	2.5	1.8	2	3.10-3.25	2.48-2.56	1.72-1.65	2.11-2.37
				(0.164)-(0.283)	(0.015)-(0.283)	(0.095)-(0.189)	(0.153)-(0.255)
α_1	α_2	Σ_{11}	Σ_{22}	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\Sigma}_{11}$	$\tilde{\Sigma}_{22}$
0.8	1.5	0.3	0.5	1.06-1.13	1.57-1.64	0.28-0.37	0.56-0.45
				(0.071)-(0.102)	(0.109)-(0.158)	(0.026)-(0.073)	(0.023)-(0.061)
r_{11}	r_{12}	r_{21}	r_{22}	$\tilde{r}_{11}$	$\tilde{r}_{12}$	$\tilde{r}_{21}$	$\tilde{r}_{22}$
45	100	100	25	44.75-52.43	100-112.75	102.35-89.64	24.72-31.02
				(0.523)-(9.372)	(1.166)-(28.11)	(01.04)-(22.05)	(2.297)-(11.20)

TABLE 5.1: Ornstein-Uhlenbeck model: Maximum likelihood estimates from 1000 simulations of model (5.2), using the exact and the -approximated transition density.

The Table 5.1 shows that, for the given sample size, the results can be correctly identified using our proposed estimation approach; even if some parameters are overestimated or underestimated, the results remain acceptable as the results belong to the confidence interval. However, we believe that these results could be further proven by using other sample sizes and by adding some of alternative assumptions for the model that we did not consider in the methodology proposed in this work, which could further complicate the estimate problem and be more time-consuming. For the random parameters, the estimates can be provided using the optimization algorithm on (3.11) using the obtained estimate results of the population parameters vector Ψ. Moreover, we conduct this simulation study by the following Figure (5.2), which shows that the empirical distribution of the most approximated estimators seems to be reasonably close to a normal distribution.

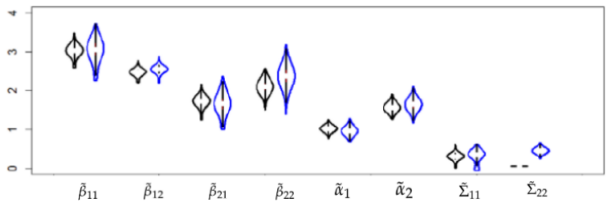

FIGURE 5.2: Empirical distribution of parameters estimates obtained using the exact and approximated transition density.

5.2 Example 2: The stochastic minimal model

The stochastic minimal model is structured by three differential equations, describing the dynamics of glucose and insulin processes in order to model the kinetics of their

measurements over time for a single individual (See Chapter 4). Here, the purpose is to estimate the stochastic minimal model to estimate the fixed and random parameters which include the insulin sensitivity as a key parameter defining the type 2 diabetes.

In this context, the approximate estimate approach proposed above is used to estimate the transition density of the continuous process in the model, that leads to obtain approximate and exact ML estimators. The estimate methodology is addressed here by simulation studies performed on healthy subjects, then a real case is achieved by using real data from the Research-Based BioPharmaceutical Company-Astrazeneca. We thus obtain a powerful and flexible modeling framework to regularize the inaccuracies of the estimation problem, often inherited from the coupled stochastic differential equations.

5.2.1 Simulation study

We start this study with an application on artificial datasets, we generate 5000 sets of simulated artificial data of dimension $3 \times (n_i + 1) \times M$ from (4.1) using the Euler-Maruyama scheme Klöden and Platen 1992 with step size of 10^{-3} and a set of the true value of parameters, that are chosen according to De Gaetano and Arino 2000; Makroglou, Li, and Kuang 2006, representing the normal range of parameters values to simulate healthy subjects (without diabetes), with M being the number of units and n_i being the number of observations or repeated measurements collected on each unit i.

For each data from 5000 generated data sets, we estimate (θ, Ψ) by applying the proposed method. We first assume that the number of repeated measurements collected on each unit is constant $n_i = m$ and $\Delta_j^i = \Delta$ for all $1 \leq i \leq M$ and $1 \leq j \leq n_i$; then we apply the methodology proposed above to estimate the SMM parameters for each dataset and we get 5000 sets of parameters estimates. We repeat this for large and small data; where we have a small number of subjects with a small number of repetitions of the experiment on each subject, as it is often the case in biomedical applications; with different possibilities of repetition of the experiment $(M, m) = (40, 60); (40, 10)$ and $(10, 20)$; for each parameter the sample mean and standard deviation are reported in tables below.

The simulation study on small data has been considered in order to check if a small

number of subjects or of the repetition of the experiment have an effect on the obtained results in using our proposed estimate methodology, and if the number of measures taken over time has a neglect effect or not, in other words, to see if it is possible to select only the essential measuring moments without repeating the measurements several times to well simulate an individual trajectory. We treat this issue in relation to our model and its study context, since in epidemiology the availability of data (measurements) at any point of time is an interesting constraint. We note that quantities of G_b and I_b are randomly simulated from the normal range of healthy subjects.

Approximate ML estimators :

To estimate the SMM parameters (θ, Ψ), we apply the methodology proposed above by incorporating (3.20) in (3.5). Then, the obtained log-likelihood function $log(L^{(a)}(\theta, \Psi))$ is approximated using Laplace approximation in order to obtain it in a closed form because the integral in the likelihood has no solution, then a numerical optimization tool is needed. So, we apply the GA and then we get the parameters estimates $\hat{\theta}$ and $\hat{\Psi}$ after choosing the right algorithm parameters (N, EN, SR, CP, MP). Finally, we obtain 5000 sets of parameters estimates. We repeat this for different possibilities of data size: $(M, m) = (40, 60)$, $(40, 10)$ and $(20, 10)$, then, we report the mean and standard deviation (std) of each parameter in Table 5.2.

True values				**(M,m)**	**Mean and (Std)**			
p_2	n	γ	h		$\hat{p}_2$	$\hat{n}$	$\hat{\gamma}$	$\hat{h}$
0.074	0.10	0.0007	90	(40,60):	0.0737 (0.0014)	0.100 (0.0013)	0.00073 ($2.02e^{-4}$)	89.08 (0.031)
				(40,10):	0.0784 (0.0059)	0.209 (0.0295)	0.00068 ($2.54e^{-4}$)	91.34 (0.712)
				(10,20):	0.0794 (0.0402)	0.156 (0.046)	0.00054 (0.0016)	62.11 (1.130)
σ_1	σ_2	σ_3	σ_{S_G}		$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_{S_G}$
0.01	0.06	0.03	0.006	(40,60):	0.009 (0.0012)	0.0616 (0.0011)	0.0343 (0.0027)	0.0061 (0.00010)
				(40,10):	0.014 (0.0023)	0.0708 (0.0060)	0.0340 (0.0035)	0.0073 (0.00030)
				(10,20):	0.015 (0.0031)	0.0938 (1.0196)	0.0480 (0.0108)	0.0067 (0.00060)
σ_{S_I}	σ_{I_0}	σ_{G_0}	μ_{S_G}		$\hat{\sigma}_{S_I}$	$\hat{\sigma}_{I_0}$	$\hat{\sigma}_{G_0}$	$\hat{\mu}_{S_G}$
0.000025	46	50	0.03	(40,60):	0.000021 ($3e^{-6}$)	44.98 (1.25)	45.96 (2.11)	0.0315 (0.0007)
				(40,10):	0.000029 ($4.4e^{-6}$)	43.94 (1.75)	45.16 (2.53)	0.0349 (0.0036)
				(10,20):	0.000016 ($0.8e^{-5}$)	41.09 (2.82)	44.64 (3.07)	0.0178 (0.0051)
μ_{S_I}	μ_{I_0}	μ_{G_0}			$\hat{\mu}_{S_I}$	$\hat{\mu}_{I_0}$	$\hat{\mu}_{G_0}$	
0.0002	95	320		(40,60):	0.00021 ($1.2e^{-6}$)	94.20 (1.12)	321.2 (1.07)	
				(40,10):	0.00025 ($1.6e^{-6}$)	92.05 (2.25)	318.6 (3.11)	
				(10,20):	0.00037 ($1.35e^{-5}$)	122.51 (2.51)	281.5 (4.88)	

TABLE 5.2: Approximated ML estimates and standard deviation from simulations of model (4.1), using the approximated transition density (3.20) for large and small DATA.

In Table 5.2, we report the approximate ML estimates obtained on large and small data by using the optimised algorithm GA. We notice that, when the sample size is large, the true values of the parameters are well identified with the exception of some in which, in all the simulations and samples, the true value does not belong to the estimated confidence interval, such as: σ_{S_I}, σ_{I_0} and σ_{G_0}; nevertheless, the results are more satisfactory when the sample size M is large for all parameters. However, the results are still satisfactory even for a sample of $M = 10$ with at least 20 measures taken on each subject. Although we can relatively accept these results for a small number of measurements, we need to know how to choose the time points to perform the measurements in the blood, as this could, physiologically, affect the observations and the selected results. Thus, it is specified here that the essential task is to know how to choose the right measurement times after the injection, even in small numbers, chosen according to the medical bases.

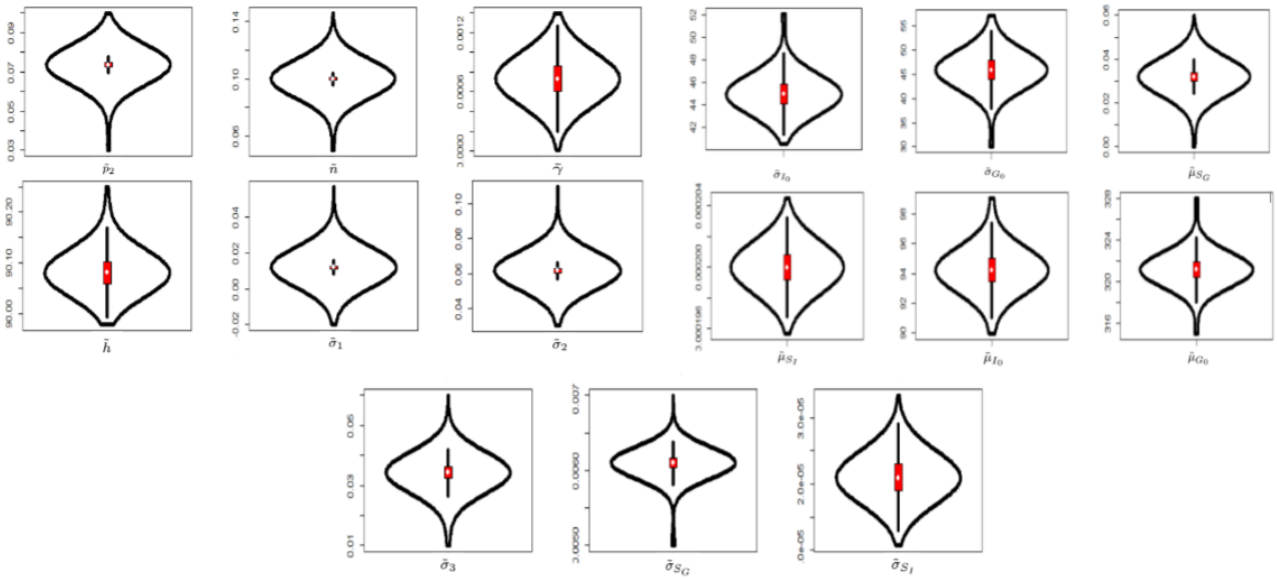

FIGURE 5.3: SMM: Empirical distribution of population parameter estimates obtained using (3.20) for (M, m) = (40,60).

The Figure (5.3) shows that, in the case of $(M; m) = (40; 60)$, the empirical distribution of the most approximated estimates seems to be reasonably close to a normal distribution.

Exact ML estimators :

Here, we estimate the transition density of the SMM using the expression (3.21) which leads to an explicit likelihood function (See Appendix A.1.2) where the integral can be solved analytically using Gauss integral properties, and then we obtain exact ML

estimators. We thus get a gradient system of fifteen equations, and then we use the Matlab function "fsolve" to obtain parameter estimates using the following appropriate initial values:

Parameters	p_2	n	h	γ	σ_1	σ_2	σ_3	μ_{S_I}
Values	0.9	0.10	90	0.00007	0.01	0.06	0.001	0
Parameters	μ_{I_0}	μ_{G_0}	σ_{S_G}	σ_{S_I}	σ_{I_0}	σ_{G_0}	μ_{S_G}	
Values	110	320	0.006	0.0000225	50	50	0.03	

TABLE 5.3: Initial values.

then we obtain 5000 sets of parameters estimates. We repeat this for different possibilities of data size: $(M, m) = (40, 60)$ and $(20, 10)$. Then, we report the mean and standard deviation (std) of each parameter in Table 5.4.

					Mean and (Standard Deviation: Std)			
True values :								
p_2	n	γ	h		$\widehat{p}_2$	$\widehat{n}$	$\widehat{\gamma}$	$\widehat{h}$
0.074	0.10	0.0007	90	M=40,m=60:	0.0799 (0.0015)	0.100 (0.0006)	0.00068 (0.00026)	90.05 (0.002)
				M=20,m=10:	0.0796 (0.0024)	0.107 (0.0134)	0.00641 (0.00048)	90.74 (0.327)
σ_1	σ_2	σ_3	σ_{S_G}		$\widehat{\sigma}_1$	$\widehat{\sigma}_2$	$\widehat{\sigma}_3$	$\widehat{\sigma}_{S_G}$
0.01	0.06	0.09	0.006	M=40,m=60:	0.015 (0.0036)	0.0600 (0.0063)	0.125 (0.0038)	0.006 (0.00023)
				M=20,m=10:	0.008 (0.0058)	0.0600 (0.0137)	0.126 (0.0198)	0.006 (0.00046)
σ_{S_I}	σ_{I_0}	σ_{G_0}	μ_{S_G}		$\widehat{\sigma}_{S_I}$	$\widehat{\sigma}_{I_0}$	$\widehat{\sigma}_{G_0}$	$\widehat{\mu}_{S_G}$
0.000025	46	50	0.03	M=40,m=60:	0.000025 ($2.1e^{-006}$)	46.08 (1.32)	10.05 (1.97)	0.030 (0.0026)
				M=20,m=10:	0.000025 ($5.2e^{-006}$)	46.14 (1.81)	5.006 (3.74)	0.030 (0.0153)
μ_{S_I}	μ_{I_0}	μ_{G_0}			$\widehat{\mu}_{S_I}$	$\widehat{\mu}_{I_0}$	$\widehat{\mu}_{G_0}$	
0.0008	95	320		M=40,m=60:	0.0009 ($0.8e^{-006}$)	95.20 (2.39)	320.0 (2.97)	
				M=20,m=10:	0.0117 ($1.3e^{-006}$)	95.72 (2.60)	319.9 (3.46)	

TABLE 5.4: Exact maximum likelihood estimates of (4.1).

The Table 5.4 shows that the parameter estimates are better identified when the size is large and most of the estimated values are close to the true values of the parameters, and while some are overestimated and others underestimated, they still belong to the estimated confidence interval. Thus, most parameters remain well identified whether in small samples, and the obtained results still acceptable. Therefore, we can conclude that the proposed estimation method appears to be suitable for pharmacokinetic/pharmacodynamic applications where data are generally scarce.

In epidemic models such as the SMM, it is important to pay attention to the appropriate timing of the measurements, as the concentration measurements change over time and

must be selected according to certain physiological criteria. We must therefore select the moments of the significant peaks, even if it is at the limit of 10 observations according to the results of our study, in order to have good interpretations. In Figure (5.4), we report the empirical distribution of the estimates, and from the vioplots below, it seems reasonable to assume that the estimates have an asymptotic normal distribution.

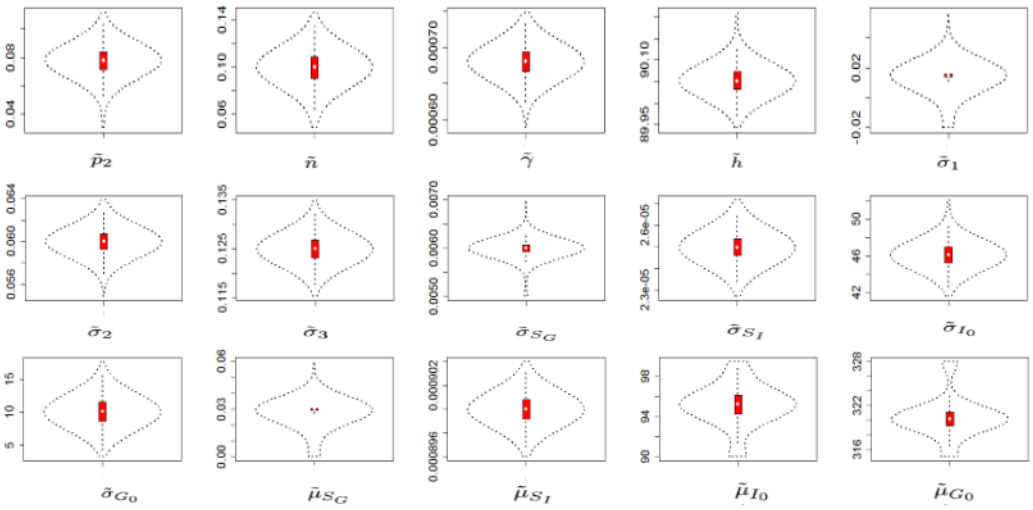

FIGURE 5.4: Empirical distribution of population parameter estimates obtained using (3.21) for (M, m) = (40,60).

In Figure (5.5) and (5.6), the boxplots of the estimates of the two key parameters of diabetes $\widehat{S}_I$ and $\widehat{S}_G$ are reported for the case of $(M, m) = (40, 60)$ and $(M, m) = (10, 20)$, respectively, where the estimates are generated from (3.11) for different units. From Figure (5.5), the empirical mean and the standard deviation (std) of the estimated random effects are, respectively: 0.0011(0.000025) and 0.03(0.006) for $\widehat{S}_I$ and $\widehat{S}_G$.

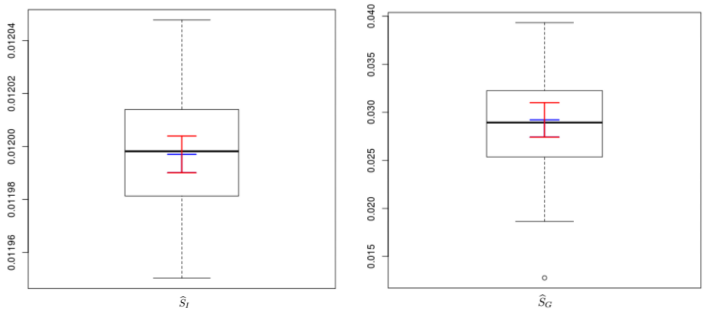

FIGURE 5.5: Boxplots of the random effects estimates of $\widehat{S}_I$ and $\widehat{S}_G$ from (3.11) for (M, m) = (40,60).

Whereas, for $(M, m) = (10, 20)$, the empirical mean and the standard deviation (std) of the estimated random effects are: 0.010(0.000025) and 0.03(0.006) for $\widehat{S}_I$ and $\widehat{S}_G$ respectively. Thus, the empirical parameters of the distribution of $\widehat{S}_I$ and $\widehat{S}_G$ in both cases are very close to their true values, which proves our simulation results.

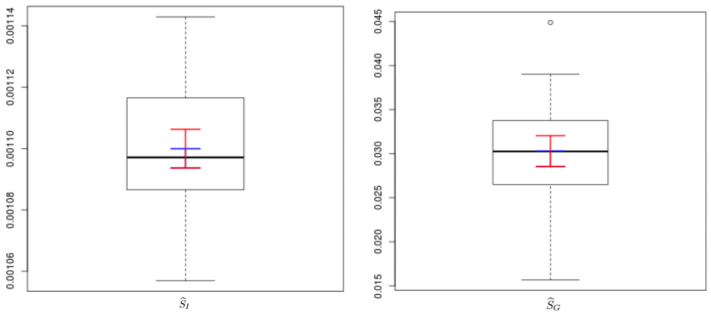

FIGURE 5.6: Boxplots of the random effects estimates of $\widehat{S}_I$ and $\widehat{S}_G$ from (3.11) for (M, m) = (10,20).

Finally, from the simulation study where we have considered two examples of SDME models, we can conclude that the parameters values of the models appear to be correctly identified using the proposed approach, based on the Risken approximation to approximate the transition density of a continuous stochastic process.

5.2.2 Real study

Real DATA Description

In this paragraph, the stochastic minimal model is estimated on real data, from the Research-Based BioPharmaceutical Company-Astrazeneca, which is obtained using an intravenous glucose tolerance test (IVGTT), where glucose and insulin concentrations in plasma are subsequently sampled after an intravenous glucose injection. The DATA report the glucose concentrations, insulin concentrations, and the C-peptide measures. The data is collected on 20 healthy subjects and 46 T2D patients. For the T2D patients, they were 15 - 34 years old with disease durations two to ten years, and classified as having T2D by the reporting physician and $Hb1Ac < 10\%$; which is the amount of glucose that sticks to the red blood cells; and for the healthy subjects they were 25-50 years old with no diabetes. Measurements of glucose and insulin concentration are measured before injection of 0.3 g /kg of the glucose and after according to the following time points for each individual: 3, 4, 5, 7, 10, 15, 20, 25, 30, 60, 115 and 120 minutes, and the $\mu U/ml$ was the measure unit.

At first, we estimate the parameters model for the healthy population and then for the T2D population, in order to compare the results of parameters model estimates describing the glucose-insulin process in two different populations (healthy and diabetic). Also, to avoid confusion about the results, the estimation procedure is performed on each population separately because they are not homogeneous and the concepts of random effects and fixed effects are then related to the same population.

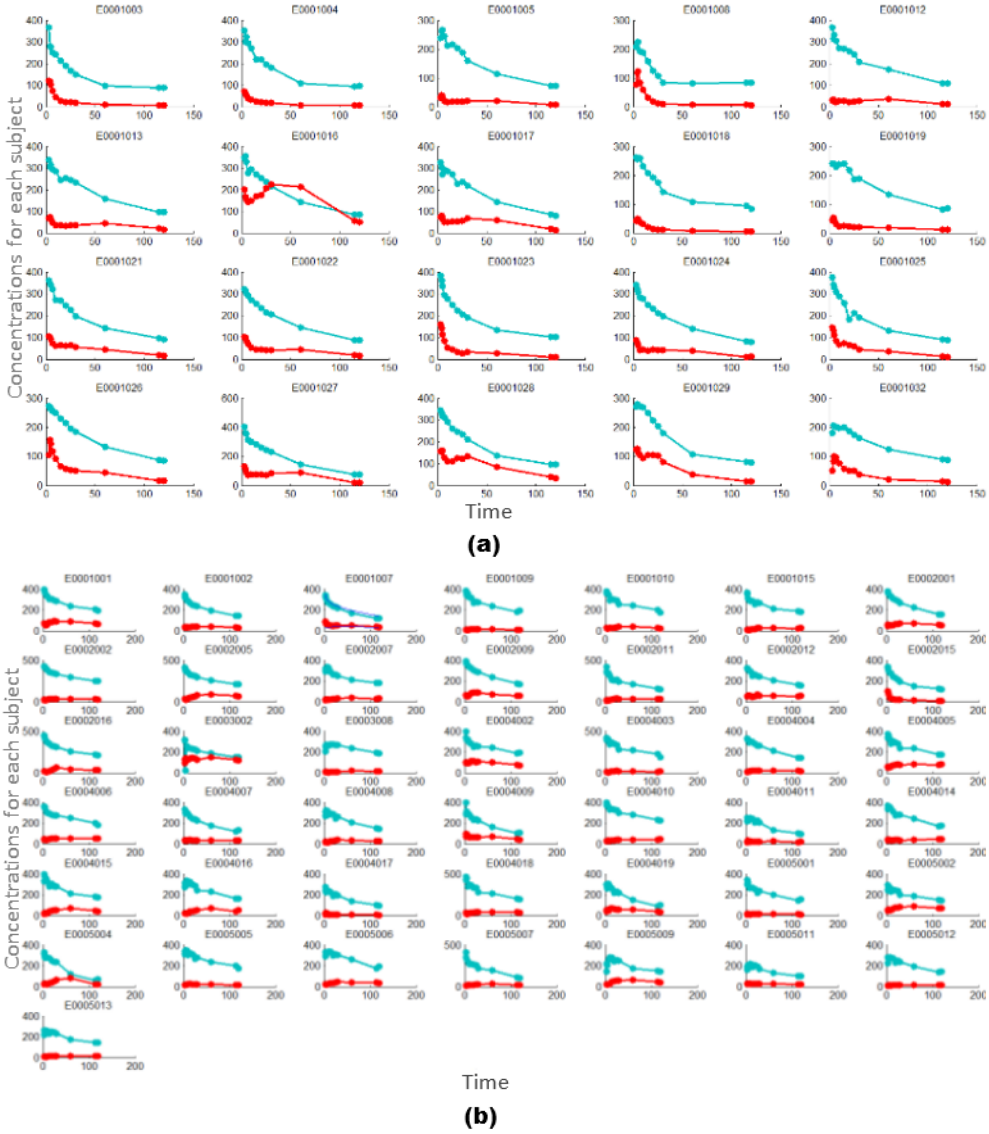

FIGURE 5.7: Plots of the insulin and glucose concentration for each healthy subject (a) and T2D patients (b), where the glucose concentration is shown in blue and the insulin concentration is shown in red $\mu U/ml$.

Augmented data

In this section, we deal with the transformation of the process in (4.1) using Lamperti transform ν: $\nu(Y) = \int^{Y_t} \frac{du}{\Sigma(u)}$ which by definition exists when the diffusion is reducible; condition for reducibility in (4.1) is obvious; then using Itô's lemma, the transformation $\mathbf{Y}_t = \nu(Y_t)$ defines a new 3-dimentional diffusion process $\mathbf{Y}_t$, being the solution of the following SDME model, where the existence and the uniqueness conditions of the solution are fulfilled (See appendix A.2):

$$d\mathbf{Y}_{\mathbf{t}}^i = \xi_L(\mathbf{Y}_t^i, t, \theta, b^i)dt + dW_t^i, \quad i = 1, ..., M, \tag{5.5}$$

where the hth element of ξ_L is given by (h = 1, . ., 3):

$$\xi_L^{(h)}(\mathbf{Y}_t^i, t, \theta, b^i) = \sum_{p=1}^{3}((\Sigma^{-1})_{hp}\xi^{(p)}(\nu^{-1}(\mathbf{Y}_t), t, \theta, b^i)) - \frac{1}{2}\sum_{p,q,s}^{3}\{\Sigma^{-1}\frac{\partial\Sigma}{\partial Y_s}(\nu^{-1}(\mathbf{Y}_t))\Sigma^{-1}\}_{hp}\Sigma_{ps}\Sigma_{qs}$$

Thus, in the following real study, the proposed estimate approach is applied on (5.5) to obtain $\hat{\theta}$ and $\hat{\Psi}$. Considering the Risken approximation proposed in Chapter 3 to approximate the transition density, we need high frequency data with a small time step, however, in our available data as in the epidimic studies, the designs are usually sparse. To overcome this constraint, we impute intermediate points between each pair of observations in order to augment the observed data $\mathbf{Y}^{obs}$, using the Eraker approach developed by Eraker 2001.

Augmented data algorithm:

The data augmentation method consists of imputing $m-1$ latent data points between each pair of observations $\mathbf{Y}^{obs}$. To ensure that discretization bias is arbitrary small we put $\Delta_t = \frac{1}{m}$ and $N = mM$ for a chosen positive integer m. Therefore [0, T] is divided into N + 1 equidistant points $t_0 = 0 < t_1 < ... < t_m < t_{m+1} < ... < t_N = T$, then the diffusion process is in state $\mathbf{Y}_{t_k}$ at time t_k which is only known on times t_j when j is an integer multiple of m and all points $\mathbf{Y}_{t_k}$, $k \neq j$ are treated as missing data. Let denotes by $\hat{\mathbf{Y}}$ the $3 \times (N+1)$ matrix obtained by stacking all elements of augmented data (observed and missing), that is :

$$\hat{\mathbf{Y}} = \begin{pmatrix} \hat{\mathbf{y}}_{1t_0} & \hat{\mathbf{y}}_{1t_1} & \cdots & \hat{\mathbf{y}}_{1t_m} & \hat{\mathbf{y}}_{1t_m+1} & \cdots & \hat{\mathbf{y}}_{1t_N} \\ \hat{\mathbf{y}}_{2t_0} & \hat{\mathbf{y}}_{2t_1} & \cdots & \hat{\mathbf{y}}_{2t_m} & \hat{\mathbf{y}}_{2t_m+1} & \cdots & \hat{\mathbf{y}}_{2t_N} \\ \hat{\mathbf{y}}_{3t_0} & \hat{\mathbf{y}}_{3t_1} & \cdots & \hat{\mathbf{y}}_{3t_m} & \hat{\mathbf{y}}_{3t_m+1} & \cdots & \hat{\mathbf{y}}_{3t_N} \end{pmatrix}.$$

Let $\hat{\mathbf{Y}}_r$ denotes the r^{th} column of $\hat{\mathbf{Y}}$ (if r is a multiple of m, $\hat{\mathbf{y}}_i$ is an observed data).

For the nonlinear SDME model (5.5), with drift and diffusion functions satisfying easily the assumptions Al-A4 in Øksendal 2003 (See Appendix A.1) the $\hat{\mathbf{Y}}_r$ can be updated using AR-MH algorithm with normal proposal density $q(./\hat{\mathbf{Y}}_{r-1}, \hat{\mathbf{Y}}_{r+1})$ with mean $\frac{\hat{\mathbf{Y}}_{r-1}+\hat{\mathbf{Y}}_{r+1}}{2}$ and variance $\frac{\Delta_t}{2}$, as is proved in Eraker 2001. So, we consider the following algorithm:

Algorithm:

1. Initialize $\hat{\mathbf{Y}}_r$, using linear interpolation between observed values of $\mathbf{Y}_r$,

2. Choose j multiple of m, $[t_j, t_{j+m}]$, the time interval of size m, in whose the path interior is to be updated,

Repeat

3. For all $r = 1, 2, ..., N-1$, at iteration h draw $\hat{\mathbf{Y}}_r^{(h)}/\hat{\mathbf{Y}}_{r-1}^{(h-1)}, \hat{\mathbf{Y}}_{r+1}^{(h-1)}, \theta, b^i)$ using the AR-MH algorithm with proposal density $\mathcal{N}(\frac{\hat{\mathbf{Y}}_{i-1}+\hat{\mathbf{Y}}_{i+1}}{2}, \frac{\Delta_t}{2})$.

4. Increase the value of h

End

Real results

Applying the proposed estimation approach for $\mathbf{Y}$, we obtain a likelihood function where the integral have no solution. Then, we use therefore the Laplace method and the optimization GA algorithm for the obtained likelihood function on real data, we finally obtain the following estimation results (mean and standard deviation (std)) in Table 5.5.

Mean and (Std) for healthy				**Mean and (Std) for patients**			
$\hat{p}_2$	$\hat{n}$	$\hat{\gamma}$	$\hat{h}$	$\hat{p}_2$	$\hat{n}$	$\hat{\gamma}$	$\hat{h}$
0.069	0.103	0.00072	91.05	402684	0.107	0.00065	90.84
(0.013)	(0.0118)	($0.1e^{-4}$)	(0.08)	(0.022)	(0.00121)	($0.3\ e^{-4}$)	(0.10)
$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_{S_G}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_{S_G}$
0.0142	0.0716	0.0385	0.005	0.013	0.079	0.0328	0.0025
(0.001)	(0.0130)	(0.00049)	(0.00032)	(0.006)	(0.012)	(0.00065)	(0.0003)
$\hat{\sigma}_{S_I}$	$\hat{\sigma}_{I_0}$	$\hat{\sigma}_{G_0}$	$\hat{\mu}_{S_G}$	$\hat{\sigma}_{S_I}$	$\hat{\sigma}_{I_0}$	$\hat{\sigma}_{G_0}$	$\hat{\mu}_{S_G}$
0.000023	44.12	45.05	0.0342	0.00002	42.013	41	0.0156
($6e^{-7}$)	(1.86)	(1.21)	(0.00053)	($0.33e^{-5}$)	(2.61)	(1.55)	(0.0007)
$\hat{\mu}_{S_I}$	$\hat{\mu}_{I_0}$	$\hat{\mu}_{G_0}$		$\hat{\mu}_{S_I}$	$\hat{\mu}_{I_0}$	$\hat{\mu}_{G_0}$	
0.0003	87.69	342.32		0.00002	41.472	380.38	
($6.3e^{-6}$)	(0.621)	(0.923)		($0.07e^{-4}$)	(1.21)	(0.78)	

TABLE 5.5: Real results.

For healthy data :The estimate value of the parameter p_2, which represents the decreased glucose absorption capacity in tissues, is weak for healthy subject, which explains why healthy individuals have a significant insulin sensitivity that is not negligible of values varying from 0.00028 to 0.000312. This considerable sensitivity to insulin allows these persons not to have a high glucose level in the blood, because it is well converted into energy in the cells and then not to become infected with Type 2 diabetes, because, a sufficient amount of insulin is produced and well used.

So, healthy people are safe from this disease thanks to a considerable insulin capability to increase glucose utilization to muscles and liver, and to have also an interesting ability of glucose to enhance its own disposal independently of the insulin level which is measured by S_G. The estimates of the insulin sensitivity S_I for 20 healthy persons belongs to: [0.00028; 0.00033], that we have obtained for each subject in the healthy population from (3.11).

For T2D patient data : We note that there is a divergence for the parameter p_2 which becomes of a great value, this can be interpreted by an interesting decrease in glucose uptake capacity in tissues for T2D patients, because of a significant decrease in insulin sensitivity for patients. These conclusions may also justify the mathematical relationship between the two parameters $S_I = p_3/p_2$. So, the value of p_2 changes from the healthy to the patient population but the parameter is not random in each population,

it changes by changing the physical characteristics of the elements constituting each population. In addition, the glucose efficiency S_G which represents the ability of glucose to be eliminated in the muscles has been decreased, so for the T2D patients, the glucose level in the blood cannot disappear easily. Also, we can note that the initial glucose concentration is higher in the T2D patients with a lower insulin concentration than in healthy subjects, this result can be explained by the fact that the T2D patient has a higher blood glucose level that is not eliminated because of the insufficient insulin secretion or of the low insulin sensitivity. From (3.11), the values of S_I for T2D individuals belongs to: [0.000018; 0.000029]; they are lower than those of healthy subjects, which shows that the T2D patients are actually less sensitive to insulin than healthy patients, which involves medical interventions to regulate their insulin levels.

5.2.3 Discussion about the stochastic minimal model

In this thesis, we adopt an estimation method as a flexible modeling framework to regularize the ill-posed problem of the stochastic minimal model, by coupling three stochastic differential equations of glucose-insulin kinetics. The processes are considered stochastics by including Brownian motion compenents to the ODEs, representing the noise of the process increments, because we believe that, physiologically, these real processes could not be deterministic. In this work, we consider all the three differential equations of the glucose and insulin concentrations in the minimal model simultaneously, the system becomes then computationally very complex. In the literature, this kind of modelling of the minimal model has been analysed in a simpler way in the late of the 1970s, where we consider the insulin part as known. However, we believe that both glucose and insulin processes constitute a single dynamic system and important information is lost when we consider the insulin as known, nevertheless, the inference statistic of the proposed stochastic minimal model is not an obvious procedure. Here, the proposed parameter estimation method is based on the statistical inference by maximizing the likelihood function of the model, for the estimation of the quantitative portrait metabolic processes for a population.

Moreover, in this thesis, the stochastic minimal model is treated using population approach, which raises inference challenges, where data from several subjects are considered simultaneously in order to study the behaviour of the entire population. So, we incorporate two types of parameters in the model, fixed effects: to represent the common parameters for all subjects and describe the common physiological side of the glucose-insulin kinetics, and random effects: which vary between individuals and

represent the inter- and intra-variability in the population. So, the obtained model is a stochastic differential mixed effects model of insulin-glucose kinetics, with an unknown transition density, that requires approximated methods in order to have the likelihood function in an approximate closed form. Therefore, we propose to derive the transition density of the process using the Fokker Planck equation and based on Risken approximation in Risken 1996, that we perform otherwise in order to get an explicit form of the likelihood function and obtain exact estimators of the parameters. Finally, the proposed approach is addressed by simulation studies on large and small data, to evaluate the obtained results for both cases, since, in the epidemic field the data are not sufficiently large and are usually sparse.

The main result of our methodology is to regularize the ill-posed estimation problem of the stochastic and coupled minimal model, by characterizing the transition density of the stochastic process via the forward or Fokker-Planck equation. Moreover, this work presents a general framework for parameter estimation of an SDME model with unknown transition density, by deriving a particular functional partial differential equation which characterizes the exact likelihood function, in order to approximate the transition density for a continuous-time stochastic process described by first-order nonlinear stochastic differential equations of the generalized ltô type. The use of continuous-time process described by ltô stochastic differential equations is now an integral part of such diverse fields as stochastic optimal control theory, financial, economics, and statistical thermodynamics. So, the proposed methodology looks like a promising framework that can be used in similar stochastic differential mixed effects models, but, we notice that although the results are acceptable using small data, it is desirable to have more large data to have more satisfactory results.

Conclusions and future directions

In this thesis, a nonlinear mixed effects model with stochastic differential equations incorporating multiple random effects, namely stochastic differential mixed effects model, has been studied; and an estimation method for such population models has been proposed and evaluated through simulations and real studies. In the applied side of the research in our thesis, the construction of a mathematical model using SDEs and the mixed effects model was tried, in order to describe the whole blood glucose-insulin system, based on Bergman's minimal model. Such a coupled stochastic model is more physiologically valid but also more complex and highly ill-posed estimation problem.

From a theoretical point of view, analysing repeated measurements data, where it is necessary to take into account the variability between experiments and the stochasticity in the individual dynamics, the SDME modelling is required in order to obtain more precise estimates of population characteristics. So, we hope that the contributing results, given in this work, be helpful to more knowledge of the stochastic mixed effects modelling, especially, in areas where mixed-effects theory is used routinely, e.g. in biomedical and pharmacokinetic/pharmacodynamic studies. Thus, we believe that such class of models provides a flexible and powerful modelling approach that will undergo increasing popularity.

Before achieving the aim, we began our thesis by presenting the preliminaries concepts that we could need for the comprehension of the theoretical and applied subjects. Then, we have explored a computationally achievable estimation approach based on the ML estimation for a nonlinear and multidimensional SDME models, by proposing an approximate closed form of the transition density of the process for such model which is usually unknown, inspired from the solution of Fokker-Planck equation proposed in Risken 1996. In this context, we have introduced all tools and algorithms necessary for the implementation of the methodology, which we can consider as a proportionately simple and flexible framework used in the estimation of this kind of models. Then, this research attempted to model the dynamics of insulin-glucose kinetics using the SDME model by giving exact estimators of key parameters of diabetes metabolism, which to our knowledge have not yet been studied in the literature.

This work has a number of limitations, mostly due to the difficulty in carrying out the closed form approximation of the likelihood for multidimensional SDME models, when the dimension of random vector parameter is rather than 2. Moreover, the proposed methodology is even more difficult when the diffusion term is not reducible, and also when it is complicated to derive the gradients and Hessians terms of the likelihood. As well as, the measurement errors are not considered in this work, which it can present a real constraint when dealing with real cases, when this noise source could not be negligible compared to the noise of the system, because, for a good stochastic version, it will be better to include noise on process increments and noise on observations that may be significant compared to system noise. However, all of these limits can provide a guideline on our future directions and perspectives, to study and generalize the proposed estimation methodology to a larger framework of application of SDME models in the study of other epidemics.

As mentioned before, insulin is secreted by the beta-cells in the portal vein and extracted by the liver before it appears in plasma. An indirect measurement approach is essential to quantify hepatic insulin extraction, as it was proposed in the thesis by using the C-peptide substance, since the direct measurement requires invasive protocols, with catheters placed in a artery and hepatic vein. So, in our future direction work, we consider that this substance is unobserved in the model which leads to consider a semi-observed SDME model, that we are interested to estimate which can give a comparable framework for the alternative approach which uses the measurements of the C-peptide.

Moreover, to improve our proposed glucose-insulin modeling and the estimation of its associated parameters, we propose a formulation of an SDME model where the drift and diffusion coefficients depend on the fixed and random parameters, as presented below in the overviwe of our future persective.

So, in our future work, we are interested in the modeling of the kinetics of glucose and insulin molecules in the organism, considered as a closed population having a total number of molecules equal to the size N. The organism is then considered to be divided into five compartments presented by organs and disposals having an important role in glucose-insulin homeostasis. Thus, the diabetes epidemic is modeled by a multidimensional SDME model where the stochastic process is a continuous-time three-dimensional Markovian jump process. In which each molecule can be found at a given moment in one of the five compartments: GS: the glucose space compartment, M:

muscles-compartment, IP: Insulin plasma compartment, IA: Remote insulin compartment, as presented in Figure (5.8). In this jump model there are four parameters of interest. Subsequently, we approximate the Markov jump process by a three-dimensional Itô diffusion process, then we describe the transitions from a state to another by a master equation, transformed after tedious calculus to a Fokker-Planck equation.

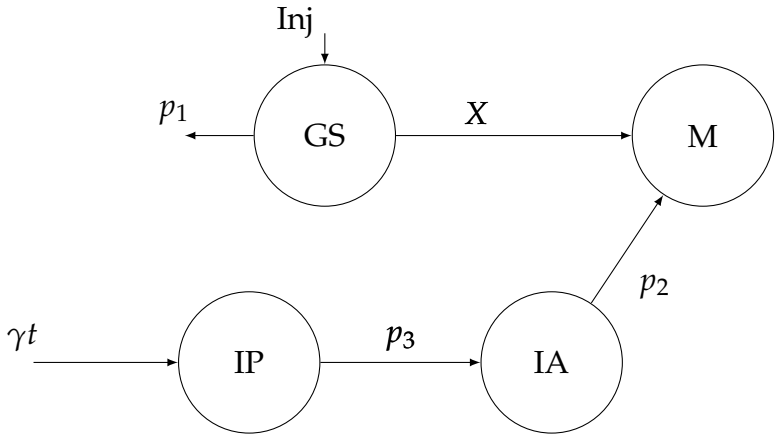

FIGURE 5.8: State diagram for glucose-inulin kinetics modelling, where GS: presents the glucose space compartment, M: muscles, IP: Insulin plasma compartment, IA: Remote insulin compartment.

Explicitly, we will assume that the time variable is continuous, $t \in [0, \infty[$ and that G(t), X(t), and I(t) are continuous random variables, that is:

$$G(t), X(t), I(t) \in [0, N]$$

where: G(t): Number of glucose molecule in plasma; X(t): Number of insulin molecule in action; I(t): Number of insulin molecule in plasma.

The process Y jumps from state $j = \begin{pmatrix} g \\ x \\ i \end{pmatrix}$ at time t to state $j + l_k$ with $l \in R^3$, $k = 1, ..7$ at time $t + \Delta t$ with transition $P_{j,j+l_k}(t)\Delta t$.

$$\begin{array}{ccc} j & p_{j,j+l_k}\Delta t & j + l_k \\ (G, X, I, M) & \longrightarrow & (G+1, X, I, M) \\ (G, X, I, M) & \longrightarrow & (G-1, X, I, M) \\ (G, X, I, M) & \longrightarrow & (G-1, X, I, M+1) \\ (G, X, I, M) & \longrightarrow & (G, X+1, I-1, M) \\ (G, X, I, M) & \longrightarrow & (G, X-1, I, M+1) \\ (G, X, I, M) & \longrightarrow & (G, X, I+1, M) \\ (G, X, I, M) & \longrightarrow & (G, X, I, M) \end{array}$$

Denote the transition pdf for the stochastic process as:

$$p(y, t + \Delta t; y', t),$$

where at time t, $Y(t) = y'$, and at time $t + \Delta t$, $Y(t + \Delta t) = y$. In the following, we consider that the pdf satisfies a forward Kolmogorov differential equation for the SDE model of glucose-insulin kinetics, which is a second order partial differential equation. The rates $p_{j,j+l_k}(t)$ are given in details in the following system which resume all jumps.

$$p_{j,j+l_k}(\Delta t) = \begin{cases} \dfrac{C.(g+1)}{N}\Delta t, & l_1 = (1,0,0), \\ \left(p_1\dfrac{(g+1)-g_b}{N} + \dfrac{x(g+1)}{N}\right)\Delta t, & l_2 = (-1,0,0), \\ p_3\dfrac{(i-1)-i_b}{N}\Delta t, & l_3 = (0,1,-1), \\ p_2\dfrac{x-1}{N}\Delta t, & l_4 = (0,-1,0), \\ \dfrac{\gamma.t(g-h)-n((i-1)-i_b)}{N}\Delta t, & l_5 = (0,0,1), \\ 1 - \dfrac{C.g + p_1(g-g_b) + xg + p_3(i-i_b) + p_2 x}{N}\Delta t & l_6 = (0,0,0), \\ + \dfrac{\gamma.t(g-h)-n(i-i_b)}{N}\Delta t, & \\ 0, & otherwise, \end{cases}$$

Let denote: $b_1(g,x,i) = \dfrac{C.g}{N}\Delta t$, $b_2(g,x,i) = \dfrac{p_1(g-g_b)+xg}{N}\Delta t$, $b_3(g,x,i) = p_3\dfrac{i-i_b}{N}\Delta t$, $b_4(g,x,i) = p_2\dfrac{x}{N}\Delta t$, and $b_5(g,x,i) = \dfrac{\gamma.t(g-g_b)-n(i-i_b)}{N}\Delta t$.
Then, the forward Kolmogorov differential equations is as follows:

$$\begin{aligned} \frac{\Delta p(g,x,i)}{\Delta t} =& p(g+1,x,i)b_1(g+1,x,i) - p(g,x,i)b_1(g,x,i) + p(g-1,x,i)b_2(g-1,x,i) \\ & - p(g,x,i)b_2(g,x,i) + p(g,x+1,i-1)b_3(g,x+1,i-1) - p(g,x,i)b_3(g,x,i) \\ & + p(g,x-1,i)b_4(g,x-1,i) - p(g,x,i)b_4(g,x,i) + p(g,x,i+1)b_5(g,x,i+1) \\ & - p(g,x,i)b_5(g,x,i). \end{aligned}$$

Now, we consider $g = \dfrac{g}{N}$, $x = \dfrac{x}{N}$ and $i = \dfrac{i}{N}$ are proportions (concentrations) of G, X and I, where N is the total number of molecules in the considered system. As $N \to \infty$, we have the following Taylor expansion approximations:

$$\frac{\partial P(g,x,i)}{\partial t} = \left[\frac{\partial}{\partial g}\left((b_1-b_2)\,(g,x,i)\right) + \frac{\partial}{\partial x}\left((b_3-b_4)\,(g,x,i)\right) + \frac{\partial}{\partial i}\left((-b_3+b_5)\,(g,x,i)\right)\right] p(g,x,i)$$
$$+\frac{1}{2}\left[\frac{\partial^2}{\partial g^2}b_2(g,x,i) + \frac{\partial^2}{\partial x^2}(b_3-b_4)(g,x,i) - 2\frac{\partial^2}{\partial x \partial i}b_3(g,x,i) + \frac{\partial^2}{\partial i^2}(b_3-b_5)(g,x,i)\right] p(g,x,i) \quad (5.6)$$

then using the arguments: $\frac{\partial}{\partial Y} = (\frac{\partial}{\partial y_k})_{y_k \in \{g,x,i\}}$ and $\frac{\partial^2}{\partial Y^2} = (\frac{\partial^2}{\partial y_k \partial y_j})_{y_k, y_j \in \{g,x,i\}}$, then the equation (5.6) can be rewritten as:

$$\frac{\partial}{\partial t}\mathbb{P}(\mathbf{y}(t)) = -\frac{\partial}{\partial \mathbf{y}}\left[\mu(\mathbf{x})\mathbb{P}(\mathbf{y},t)\right] + \frac{1}{2}\frac{\partial^2}{\partial \mathbf{y}^2}\left[\mathbf{\Sigma}(\mathbf{y})\mathbb{P}(\mathbf{y},t)\right], \quad (5.7)$$

where:

$$\mu(Y_t,t,\theta,b^i) = \begin{pmatrix} c - p_1(g-g_b) + xg \\ p_3(i-i_b) - p_2 x \\ -n(i-i_b) + \gamma(g-h)t - p_3(i-i_b) \end{pmatrix}$$

$$\Sigma(Y_t,t,\theta,b^i) = \begin{pmatrix} p_1(g-g_b) - xg & 0 & 0 \\ 0 & p_3(i-i_b) + p_2 x & -p_3(i-i_b) \\ 0 & -p_3(i-i_b) & p_3(i-i_b) - n(i-i_b) + \gamma(g-h)t \end{pmatrix}$$

The equation (5.7) is the Fokker-Plank equation associated to a diffusion process $\mathbf{Y} = (Y_t,\ t \geq 0)$ which is solution, according to Øksendal 2003, chap.7, to the following non-linear three-dimensional Itô stochastic differential equation:

$$dY_t = \begin{pmatrix} -p_1(g-g_b) + xg \\ p_3(i-i_b) - p_2 x \\ -n(i-i_b) + \gamma(g-h)t - p_3(i-i_b) \end{pmatrix} dt + \boldsymbol{\sigma}(Y_t,b,\theta,t) \begin{pmatrix} dW_1 \\ dW_2 \\ dW_3 \end{pmatrix}$$

So, that:

$$dY_t = \mu(Y_t,\theta,b,t)dt + \boldsymbol{\sigma}(Y_t,\theta,b,t)dW(t), \quad (5.8)$$

where $\boldsymbol{\sigma}$ is such that $\boldsymbol{\sigma}\boldsymbol{\sigma}^T = \boldsymbol{\Sigma}$, here $\boldsymbol{\sigma}^T$ is transposed matrix of the matrix $\boldsymbol{\sigma}$. The SDME model obtained in (5.8) is more realistic where variations are incorporated into the diffusion term through the variables and parameters on which the term depends. So, by using the Bayesian inference to explore discrete data using a diffusion approximation, what estimation constraint can we have when the model is semi-observed? what degree of similarity can we have between the estimated measures of X and the measures used previously? Can the transition density approximation based on the Risken approximation, previously used, lead to effective results in this case?

.1 Appendix A

.1.1 Verification of assumptions

A.1

The uniqueness and existence of a strong solution of the equation (4.1), with a constant diffusion term, are guaranteed by two conditions, local Lipschitz and linear growth of both coefficients drift and diffusion, see Liptser and Shiryaev 2001; Kutoyants 2004 for details. To verify the local Lipschitz for ξ, let Y_1 and Y_2 be in $\mathbb{R}^3$, we have after some simple calculus:

$$\| \xi(Y_1) - \xi(Y_2) \| \leq K \| Y_1 - Y_2 \|$$

where $K = max K_1, K_2, K_3$ and $K_1 = S_G + 1$, $K_2 = p_2$ and $K_3 = n$. for the linear growth of ξ coefficient we have after simple calculus:

$$\| \xi(Y, \theta, b) \|^2 \leq C(1 + \| Y \|^2)$$

where $C = max C_1, C_2, C_3$ and $C_1 = S_G^2$, $C_2 = p_2^2$ and $C_3 = n^2$.

A.2

With the same way, the uniqueness and existence of a strong solution of the obtained equation in (5.5) are guaranteed by local Lipschitz and linear growth conditions of the transformed coefficients drift ξ_L. let $\mathbf{Y}_1$ and $\mathbf{Y}_2$ be in $\mathbb{R}^3$, we have after simple calculus:

$$\| \xi_L(\mathbf{Y}_1) - \xi_L(\mathbf{Y}_2) \| \leq M \| \mathbf{Y}_1 - \mathbf{Y}_2 \|$$

and

$$\| \xi_L(\mathbf{Y}, \theta, b) \| \leq S(1 + \| \mathbf{Y} \|^2)$$

where : $M = maxM_1, M_2, M_3$ and $M_1 = \frac{S_G + 1}{\sigma_1} + 1$, $M_2 = \frac{p_2}{\sigma_2}$, $M_3 = \frac{n}{\sigma_3}$, and $S = maxS_1, S_2, S_3$ with $S_1 = \frac{S_G^2}{\sigma_1^2}$, $S_2 = \frac{p_2^2}{\sigma_2^2}$ and $S_3 = \frac{n^2}{\sigma_3^2}$

.1.2 Proof of (3.21)

$$\begin{aligned}
P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) &= [1 - \frac{\partial}{\partial Y_j}\mu(Y_j, t_j, \theta, b)\Delta_j + \frac{\partial^2}{\partial {Y_j}^2}\Sigma(Y_j, t_j, \theta, b)\Delta_j]\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iu(y_j - y_{j-1})}du \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}[e^{iu(y_j - y_{j-1})} - iu\mu(Y_j, t_j, \theta, b)e^{iu(y_j - y_{j-1})}\Delta_j - \frac{\partial}{\partial Y_j}\mu(Y_j, t_j, \theta, b)e^{iu(y_j - y_{j-1})} \\
&\quad -u^2\Sigma(Y_j, t_j, \theta, b)e^{iu(y_j - y_{j-1})}\Delta_j + \frac{\partial^2}{\partial {Y_j}^2}\Sigma(Y_j, t_j, \theta, b)e^{iu(y_j - y_{j-1})}\Delta_j]du \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} exp(-iu\mu(Y_j, t_j, \theta, b)\Delta_j - \frac{\partial}{\partial Y_j}\mu(Y_j, t_j, \theta, b)\Delta_j - u^2\Sigma(Y_j, t_j, \theta, b)\Delta_j \\
&\quad + \frac{\partial^2}{\partial {Y_j}^2}\Sigma(Y_j, t_j, \theta, b)\Delta_j + iu(y_j - y_{j-1}))du
\end{aligned}$$

So, using the Gaussian integral, we obtain:

$$\begin{aligned}
P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) &= \frac{1}{2\sqrt{\pi\Sigma(Y_j, t_j, \theta, b)\Delta_j}} exp(\frac{\partial}{\partial Y_j}\mu(Y_j, t_j, \theta, b)\Delta_j + \frac{\partial^2}{\partial {Y_j}^2}\Sigma(Y_j, t_j, \theta, b)\Delta_j \\
&\quad \frac{[y_j - y_{j-1} - \mu(Y_j, t_j, \theta, b)\Delta_j]^2}{4\Sigma(Y_j, t_j, \theta, b)\Delta_j})
\end{aligned}$$

Then, we extend the formula for N variables:

$$\begin{aligned}
P_Y(y_j, \Delta_j | y_{j-1}, b, \theta) &= (2\sqrt{\Pi\Delta_j})^{-N}[Det\Sigma]^{-\frac{1}{2}} exp(\frac{\partial}{\partial Y_{j_l}}\mu_l(Y_j, t_j, \theta, b)\Delta_j + \frac{\partial^2}{\partial {Y_{j_l}}^2}\Sigma_l(Y_j, t_j, \theta, b)\Delta_j \\
&\quad -\frac{1}{4\Delta_j}[\Sigma(Y_j, t_j, \theta, b)^{-1}]_{lk}[y_{j_l} - {y_{j-1}}_l - \mu(Y_j, t_j, \theta, b)\Delta_j][y_{j_k} - {y_{j-1}}_k - \mu(Y_j, t_j, \theta,
\end{aligned}$$

.1.3 Details

Likelihood function of (4.1) using (3.21):

$$\ln \mathcal{L}^{(a)}(\theta,\Psi) = \sum_{i=1}^{M}(-2\ln(2\pi) - \frac{1}{2}\ln(\sigma_{S_G}\sigma_{S_I}\sigma_{I_0}\sigma_{G_0}) + \ln(\sum_{j=1}^{n_i}(2\sqrt{\pi\Delta_j})^{-3n_i}(\sigma_1\sigma_2\sigma_3)^{\frac{-n_i}{2}} exp(-(p_2 + n)n_i)) + \ln(Cte1) + \frac{1}{2}\ln(\frac{\pi}{a_{10}}) + \frac{b_1^2}{4a_{10}} + \frac{1}{2}\ln(\frac{\pi}{a_2}) + CC_1 + \frac{1}{2}\ln(\frac{\pi}{E_3}) + E_1 + \frac{E_2^2}{4E_3})$$

The fifteen Gradients of the likeliood function for the model (4.1) using (3.21)

$$GRad_{p_2} = \sum_{i=1}^{M}(-4n_iE_3^2 - 2O_1A_2E_3 + 4E_3^2O_2A_1 + 8E_2E_3(O_2A_2 - O_1A_1) - 4E_2^2O_1A_2) = 0$$

$$GRad_n = \sum_{i=1}^{M}(nn_i + F_1 + nF_2) = 0$$

$$GRad_\gamma = \sum_{i=1}^{M}(\gamma(l_4 - \frac{1}{2\Delta_1}\frac{1}{\sigma_3}l_2^2 + \frac{2l_{22}^2}{(4a_2)^2} + l_3 - \frac{1}{4\Delta_1}\frac{1}{\sigma_3}(2l_{11} + 2I_1l_{12}) + \frac{2l_{22}l_{21}}{(4a_2)^2}) = 0$$

$$GRad_h = \sum_{i=1}^{M}(h(g_{12} + g_{24} + \frac{1}{2a_2}g_{14}^2) + g_{11} + g_{23} + \frac{1}{2a_2}g_{14}g_{13}) = 0$$

$$GRad_{\mu_{G_0}} = \sum_{i=1}^{M}(\mu_{G_0}(\frac{1}{2a}\frac{1}{\sigma_{G_0}^2} - \frac{1}{\sigma_{G_0}}) + \frac{1}{2a}C_7^1\frac{1}{\sigma_{G_0}}) = 0$$

$$GRad_{\mu_{S_G}} = \sum_{i=1}^{M}(\mu_{S_G}(\frac{1}{\sigma_{S_G}} - 1) - b_1^1) = 0$$

$$GRad_{\mu_{S_I}} = \sum_{i=1}^{M}(\mu_{S_I}(\frac{1}{2E_3} - 1) + \frac{E_{12}2E_3\sigma_{S_I}}{)} = 0$$

$$GRad_{\mu_{I_0}} = \mu_{I_0}(2Ma_2\sigma_{I_0} - 1) - b_{12}\sigma_{I_0} = 0$$

$$GRad_{\sigma_3} = \sum_{i=1}^{M}(\frac{-n_i}{2\sigma_3} - \frac{D}{\sigma_3} + \frac{B}{4\Delta_1\sigma_3^2} + \frac{\frac{2b_2C}{4\Delta_1\sigma_3^2} + \frac{b_2^2}{\Delta_1\sigma_3^2}}{(4a_2)^2} + \frac{1}{8a_2\Delta_1\sigma_3^2}) = 0$$

$$GRad_{\sigma_2} = \sum_{i=1}^{M}(\frac{-n_i}{2\sigma_2} - \frac{\sum_{j=1}^{n_i}\frac{-A_2^2}{8\Delta_j\sigma_2^2}}{E_3} + \sum_{j=1}^{n_i}\frac{A_1^2}{4\Delta_j\sigma_2^2} + \frac{8E_2E_3\sum_{j=1}^{n_i}\frac{A_1A_2}{2\Delta_j\sigma_2^2} - 4E_2^2\sum_{j=1}^{n_i}\frac{-A_2^2}{4\Delta_j\sigma_2^2}}{16E_3^2} = 0$$

$$GRad_{\sigma_{S_G}} = \sum_{i=1}^{M}(\frac{-1}{2\sigma_{S_G}} + \frac{\mu_{S_G}^2}{2\sigma_{S_G}^2} + \frac{1}{4a_1\sigma_{S_G}^2} + \frac{-2b_1 + 2b_1^2\mu_{S_G}}{16a_1^2\sigma_{S_G}^2}) = 0$$

$$GRad_{\sigma_{I_0}} = \sum_{i=1}^{M}(\frac{-1}{2\sigma_{I_0}} + \frac{\mu_{I_0}^2}{2\sigma_{I_0}^2} + \frac{-8b_2a_2\mu_{I_0} + 2b_2}{16a_2^2\sigma_{I_0}^2} - \frac{-1}{4a_2\sigma_{I_0}^2}) = 0$$

$$GRad_{\sigma_{S_I}} = \sum_{i=1}^{M}(\sigma_{S_I} - \mu_{S_I}^2 + \frac{8\mu_{S_I}E_2E_3 + 2E_2^2}{8E_3^2} + \frac{1}{2E_3}) = 0$$

$$GRad_{\sigma_1} = \sum_{i=1}^{M}(\frac{-n_i}{2\sigma_1} + \sum_{j=1}^{n_i}\frac{C_5^2}{4\Delta_j\sigma_1^2} + \frac{C_7C_1 + 2C_7^2}{32a^2\Delta_1\sigma_1^2} + \frac{1}{8a\Delta_1\sigma_1^2} - \frac{\sum_{j=1}^{n_i}\frac{-C_6^2}{4\Delta_j\sigma_1^2} - \frac{4C_2\Delta_1}{(2\sigma_{G_0}+4\Delta_1\sigma_1)^2}}{2a_{10}} + \frac{8a_{10}b_{10}(\sum_{j=1}^{n_i}\frac{C_5C_6}{2\Delta_j\sigma_1^2} + \frac{C_3}{4\Delta_1\sigma_1^2} + \frac{4C_4\Delta_1\sigma_{G_0}}{(2\sigma_{G_0}+4\Delta_1\sigma_1)^2}) - 4b_{10}^2(\sum_{j=1}^{n_i}\frac{-C_6^2}{4\Delta_j\sigma_1^2} - \frac{4C_2\Delta_1}{(2\sigma_{G_0}+4\Delta_1\sigma_1)^2})}{16a_{10}^2} = 0$$

$$GRad_{\sigma_{G_0}} = \sum_{i=1}^{M}(\frac{-1}{2\sigma_{G_0}} + \frac{\mu_{G_0}^2}{2\sigma_{G_0}^2} + \frac{-4C_7\mu_{G_0} + 4C_7^2}{32a^2\sigma_{G_0}^2} + \frac{1}{4a\sigma_{G_0}^2} + \frac{C_2}{a_{10}(2\sigma_{G_0} + 4\Delta_1\sigma_1)^2} + \frac{8a_{10}b_{10}(C_4(2\sigma_{G_0} + 4\Delta_1\sigma_1) - 2C_4\sigma_{G_0}) + 8b_{10}^2C_2}{16a_{10}^2(2\sigma_{G_0} + 4\Delta_1\sigma_1)^2}) = 0$$

where:

$$l_1 = I_j - I_{j-1} + n(I_j - I_b)\Delta_j$$

$$l_2 = (G_j - h)t_j\Delta_j$$

$$l_3 = \sum_{j=2}\frac{-2l_1l_2}{4\Delta_j\sigma_3}$$

$$l_4 = \sum_{j=2}\frac{2l_2^2}{4\Delta_j\sigma_3} = 0$$

$$l_{11} = n(I_1 - I_b)\Delta_1 \quad l_{12} = -(G_1 - h)t_1\Delta_1 \quad l_{21} = \frac{2I_1 + 2l_{11}}{4\Delta_1\sigma_3} + \frac{\mu_{I_0}}{\sigma_{I_0}} \quad l_{22} = \frac{2l_{12}}{4\Delta_1\sigma_3}$$

$$g_1 = l_1 - \gamma G_jt_j\Delta_j \quad g_2 = \gamma t_j\Delta_j \quad g_{11} = \sum_{j=2}\frac{-2(g_1 + g_2)}{4\Delta_j\sigma_3} \quad g_{12} = \sum_{j=2}\frac{-2g_2^2}{4\Delta_j\sigma_3} \quad g_3 = n(I_1 - I_b)\Delta_1 - \gamma G_1t_1\Delta_1 \quad g_4 = \gamma t_1\Delta_1 \quad g_{23} = \frac{-2(I_1 + g_3)g_4}{4\Delta_1\sigma_3} \quad g_{24} = \frac{-2g_4^2}{4\Delta_1\sigma_3} \quad g_{13} = \frac{2I_1 - g_3 - \frac{\mu_{I_0}}{\mu_{G_0}}}{4\Delta_1\sigma_3} \quad g_{14} = \frac{2g_4}{4\Delta_1\sigma_3}$$

$$A = [n(I_1 - I_b) - \gamma(G_1 - h)t_1]\Delta_1 \quad C = -2(I_1 + A) \quad B = I_1^2 + A^2 + 2I_1A \quad B = C^2/4 \quad D = \sum_{j=2}\frac{-A_{3j}^2}{4\Delta_j\sigma_3}$$

$$a_2 = \frac{1}{4\Delta_1\sigma_3} + \frac{1}{2\sigma_{I_0}} \quad b_2 = \frac{-C}{4\Delta_1\sigma_3} + \frac{\mu_{I_0}}{\sigma_{I_0}}$$

$$CC1 = D - \frac{B}{4\Delta_1\sigma_3} - \frac{\mu_{I_0}^2}{2\sigma_{I_0}} + \frac{b_2^2}{4a_2}$$

$$A_1 = X_j - X_{j-1} + p_2 X_j \Delta_j \quad A_2 = p_2(I_j - I_b)\Delta_j$$

$$E_1 = \sum_{j=1} \frac{-A_1^2}{4\Delta_j\sigma_2} - \frac{\mu_{S_I}^2}{2\sigma_{S_I}} \quad E_2 = \sum_{j=1} \frac{-A_1 A_2}{2\Delta_j\sigma_2} + \frac{\mu_{S_I}}{\sigma_{S_I}} \quad E_3 = \sum_{j=1} \frac{A_2^2}{4\Delta_j\sigma_2} + \frac{1}{2\sigma_{S_I}}$$

$$O_1 = \sum_{j=1} \frac{(I_j - I_b)\Delta_j}{2\Delta_j\sigma_2} \quad O_2 = \sum_{j=1} \frac{-X_j\Delta_j}{2\Delta_j\sigma_2}$$

$$f_1 = \sum_{j=2} \frac{-(I_j - I_b)\Delta_j(I_j - I_{j-1} - \gamma(G_j - h)t_j\Delta_j)}{2\Delta_j\sigma_3}$$

$$f_2 = \sum_{j=2} \frac{(I_j - I_b){\Delta_j}^2}{2\Delta_j\sigma_3}$$

$$f_3 = 2(I_1 - I_b){\Delta_1}^2 \quad f_4 = 2I_1(I_1 - I_b)\Delta_1 - 2(I_1 - I_b)\Delta_1^2\gamma(G_1 - h)t_1$$

$$f_5 = \frac{(I_1 - I_b)\Delta_1}{2\Delta_1\sigma_3}$$

$$C_1 = -2G_1 - 2X_1G_1\Delta_1 \quad C_2 = \frac{C_4^2}{4} \quad C_3 = 2(G_1 - G_b)\Delta_1^2X_1G_1 + 2G_1(G_1 - G_b)\Delta_1 \quad C_4 = -2(G_1 - G_b)\Delta_1 \quad C_5 = G_j - G_{j-1} + X_jG_j\Delta_j \quad C_6 = (G_j - G_b)\Delta_j$$

$$a = \frac{1}{4\Delta_1\sigma_{11}} + \frac{1}{2\sigma_{G_0}} \quad C_7 = \frac{-C_1}{4\Delta_1\sigma_{11}} + \frac{\mu_{G_0}}{\sigma_{G_0}}$$

$$cte1 = exp(\sum_{j=2} \frac{-C_5^2}{4\Delta_j\sigma_1} - \frac{\mu_{S_G}^2}{2\sigma_{S_G}} - frac\mu_{G_0}^2 2\sigma_{G_0} + \frac{C_7^2}{4a})\sqrt{\frac{\pi}{8}}$$

$$a_{10} = \sum_{j=2} \frac{C_6^2}{4\Delta_j\sigma_1} + \frac{1}{2\sigma_{S_G}} + \frac{4C_2\Delta_1}{2\sigma_{G_0} + 4\Delta_1\sigma_1}$$

$$b_{10} = \sum_{j=2} \frac{-C_5C_6}{2\Delta_j\sigma_1} + \frac{\mu_{S_G}}{\sigma_{S_G}} - \frac{C_3}{4\Delta_1\sigma_1} - \frac{C_4\sigma_{G_0}}{2\sigma_{G_0} + 4\Delta_1\sigma_1}$$

.2 Real DATA

G	X	I	time
20,3	2,0191E-06	8,40346E-07	3
15,6	2,3832E-06	8,05621E-07	4
15,3	2,2177E-06	7,08391E-07	5
13,9	1,8867E-06	5,06986E-07	7
13,4	1,6219E-06	3,1947E-07	10
11,9	1,5226E-06	1,9446E-07	15
10,5	0,000001324	1,5279E-07	20
9,3	1,2578E-06	1,5279E-07	25
8,3	1,1916E-06	1,31955E-07	30
5,4	7,944E-07	6,25051E-08	60
4,9	4,965E-07	4,86151E-08	115
4,9	5,296E-07	4,86151E-08	120
19,5	1,7543E-06	4,86151E-07	3
16,7	1,5557E-06	4,09061E-07	4
17,9	1,6219E-06	3,64613E-07	5
16,3	1,5557E-06	2,76411E-07	7
15,1	1,5888E-06	2,29185E-07	10
12,2	1,4233E-06	1,76403E-07	15
12,1	1,3902E-06	1,5279E-07	20
10,9	0,000001324	1,31955E-07	25
10	1,4233E-06	1,34039E-07	30
6	1,3571E-06	5,97271E-08	60
5,3	8,937E-07	4,79206E-08	115
5,4	7,944E-07	4,79206E-08	120
13,2	0,000000993	2,20157E-07	3
14,5	1,0923E-06	2,81273E-07	4
14,8	1,0261E-06	2,31269E-07	5

13,6	9,599E-07	1,45151E-07	7
11,8	8,937E-07	1,19454E-07	10
12	8,937E-07	1,29872E-07	15
11,3	9,268E-07	1,21538E-07	20
10,5	9,599E-07	1,3265E-07	25
8,9	1,0592E-06	1,40984E-07	30
6,3	1,1254E-06	1,40984E-07	60
4	6,289E-07	5,55601E-08	115
4	6,289E-07	5,90326E-08	120
11,5	2,0522E-06	5,31293E-07	3
12,4	2,4494E-06	8,2854E-07	4
12,5	2,4494E-06	8,55625E-07	5
10,7	2,1846E-06	5,83381E-07	7
10,5	1,9198E-06	4,08367E-07	10
8,7	1,7212E-06	2,25018E-07	15
6,9	0,000001655	1,29872E-07	20
6	1,3902E-06	8,26456E-08	25
4,6	1,0261E-06	6,18106E-08	30
4,5	8,606E-07	5,97271E-08	60
4,6	7,613E-07	4,93096E-08	115
4,6	7,282E-07	4,30591E-08	120
20,4	1,3902E-06	2,14601E-07	3
17,4	1,3902E-06	2,30574E-07	4
18,5	1,2578E-06	1,64597E-07	5
17	1,4233E-06	1,59735E-07	7
15	1,4564E-06	1,84043E-07	10
14,9	1,4564E-06	1,82654E-07	15
14,3	1,4233E-06	1,56957E-07	20

13,5	1,5557E-06	1,65291E-07	25
11,5	1,6219E-06	1,96544E-07	30
9,6	2,0191E-06	2,5002E-07	60
6,1	1,5557E-06	9,72301E-08	115
6,1	1,5226E-06	9,93136E-08	120
18,7	2,2177E-06	4,68094E-07	3
17	1,8867E-06	5,20181E-07	4
17,4	2,0522E-06	4,43092E-07	5
16,3	1,7874E-06	3,32666E-07	7
15,9	1,8205E-06	2,51409E-07	10
13,6	1,7212E-06	2,60438E-07	15
14,1	1,8205E-06	2,36825E-07	20
13,6	1,8867E-06	2,58354E-07	25
12,9	1,9529E-06	2,54187E-07	30
8,7	2,4494E-06	3,06275E-07	60
5,4	1,7212E-06	1,45845E-07	115
5,4	1,6219E-06	1,09037E-07	120
19,3	3,3431E-06	1,4022E-06	3
19,8	2,8466E-06	1,15357E-06	4
18,2	2,2508E-06	1,08203E-06	5
15,4	2,3832E-06	9,91053E-07	7
16,4	2,4163E-06	1,04106E-06	10
15	2,7142E-06	1,16954E-06	15
14,1	2,7804E-06	1,22718E-06	20
13,1	3,1776E-06	1,44803E-06	25
12	3,2107E-06	1,56679E-06	30
8	0,00000331	1,48623E-06	60
4,8	1,5226E-06	3,91004E-07	115

4,7	1,4564E-06	3,57668E-07	120
18,1	0,000002317	5,11153E-07	3
16,9	2,2508E-06	5,50739E-07	4
15,1	2,2508E-06	4,50731E-07	5
16,2	2,1515E-06	3,50723E-07	7
15,8	2,1515E-06	3,53501E-07	10
15,1	2,2839E-06	3,71558E-07	15
12,6	2,3832E-06	3,72253E-07	20
13,2	2,5818E-06	3,8267E-07	25
12,1	2,7804E-06	4,75039E-07	30
7,9	2,8135E-06	4,09061E-07	60
4,7	0,000001986	1,30566E-07	115
4,5	1,6219E-06	9,09796E-08	120
14,6	0,000001324	2,9933E-07	3
14,2	1,4895E-06	3,42389E-07	4
14,3	1,4564E-06	3,41E-07	5
14,4	1,5226E-06	2,63216E-07	7
12,9	1,5888E-06	2,18073E-07	10
11,5	1,5226E-06	1,50012E-07	15
10,7	0,000001324	9,93136E-08	20
9,7	1,2247E-06	8,95906E-08	25
7,9	1,1254E-06	7,77841E-08	30
6	0,000000993	4,93096E-08	60
5,3	0,000000662	3,95866E-08	115
4,7	0,000000662	3,81976E-08	120
13,4	1,5888E-06	3,16692E-07	3
13,3	1,5226E-06	3,68086E-07	4
13,2	1,3571E-06	3,11831E-07	5

12,7	1,3902E-06	2,26407E-07	7
13,2	1,2909E-06	1,57652E-07	10
13,3	1,3902E-06	1,67375E-07	15
12,1	1,3902E-06	1,6043E-07	20
10,3	1,3571E-06	1,42373E-07	25
10,4	1,4233E-06	1,36817E-07	30
7,4	1,5888E-06	1,27788E-07	60
4,5	1,1585E-06	8,05621E-08	115
4,8	1,1916E-06	8,33401E-08	120
20	2,2508E-06	7,36866E-07	3
19,3	2,4163E-06	6,93806E-07	4
19	2,2839E-06	6,49358E-07	5
17,8	2,2508E-06	5,18792E-07	7
15,1	2,2508E-06	4,36841E-07	10
14,9	2,0853E-06	4,54204E-07	15
13,7	2,1184E-06	4,22257E-07	20
12,6	2,1515E-06	4,49342E-07	25
10,9	2,4163E-06	3,88921E-07	30
7,9	2,2508E-06	3,12525E-07	60
5,3	1,4564E-06	1,20843E-07	115
5,1	1,4233E-06	1,12509E-07	120
17,9	0,000002317	7,16031E-07	3
17,3	2,2177E-06	6,81305E-07	4
16,9	2,2839E-06	5,96576E-07	5
16,3	2,1184E-06	4,89623E-07	7
15,1	2,0191E-06	3,75031E-07	10
14,2	1,9198E-06	3,11136E-07	15
13	1,9529E-06	3,13914E-07	20

11,9	2,0522E-06	2,90301E-07	25
11,5	2,0522E-06	2,82662E-07	30
8,1	2,6149E-06	3,07664E-07	60
4,9	1,6219E-06	1,2501E-07	115
4,9	1,5226E-06	1,02786E-07	120
21,4	3,0783E-06	1,1112E-06	3
20,1	2,9128E-06	9,86191E-07	4
18,6	2,6149E-06	7,84092E-07	5
16,4	2,2508E-06	5,79908E-07	7
15,5	1,8536E-06	3,64613E-07	10
13,8	1,8205E-06	3,00719E-07	15
12,4	1,5888E-06	2,36825E-07	20
11,5	1,5226E-06	1,85432E-07	25
10,7	1,6881E-06	2,20851E-07	30
7,5	1,6219E-06	1,90293E-07	60
5,7	1,0261E-06	7,50061E-08	115
5,6	9,268E-07	7,36171E-08	120
19	2,3501E-06	6,14633E-07	3
17,8	2,2177E-06	5,19487E-07	4
17,1	1,8867E-06	4,49342E-07	5
15,8	1,7543E-06	2,88218E-07	7
15,4	1,9529E-06	3,15998E-07	10
13,8	2,0191E-06	2,78495E-07	15
12,9	1,9529E-06	3,10442E-07	20
11,8	2,1184E-06	2,97246E-07	25
11	2,2508E-06	2,94468E-07	30
7,7	2,4494E-06	2,66688E-07	60
4,6	1,4233E-06	7,29226E-08	115

4,4	1,4895E-06	8,75071E-08	120
20,8	2,7142E-06	9,99387E-07	3
18,9	2,6811E-06	9,08407E-07	4
18,3	2,5156E-06	7,62562E-07	5
17,2	2,3832E-06	5,83381E-07	7
16	2,1846E-06	4,69483E-07	10
14,3	2,3501E-06	5,11153E-07	15
10,1	0,000002317	4,45175E-07	20
11,7	2,1846E-06	4,07672E-07	25
10,7	2,2177E-06	3,09747E-07	30
7,2	2,2508E-06	2,5766E-07	60
5	1,0923E-06	8,05621E-08	115
4,9	1,0592E-06	7,63951E-08	120
15,2	0,000002648	7,20892E-07	3
14,7	3,0121E-06	1,08759E-06	4
14,8	3,0783E-06	9,91747E-07	5
14,2	2,8466E-06	8,10483E-07	7
13,9	2,5487E-06	6,3269E-07	10
12,8	2,4163E-06	4,62538E-07	15
11,9	2,3501E-06	3,97949E-07	20
10,9	2,2839E-06	3,65308E-07	25
10,2	2,4494E-06	3,47251E-07	30
7,4	2,4825E-06	3,03497E-07	60
4,8	1,7212E-06	1,1876E-07	115
4,7	1,5226E-06	1,05564E-07	120
22,5	2,3501E-06	9,16741E-07	3
20,1	2,3501E-06	7,57006E-07	4
19,7	2,2177E-06	6,29218E-07	5

17,5	2,0191E-06	4,9379E-07	7
16,6	1,8205E-06	5,13236E-07	10
15,8	2,0191E-06	5,13931E-07	15
14,4	2,1846E-06	5,18098E-07	20
13,6	2,2839E-06	5,00735E-07	25
12,9	2,5487E-06	5,7088E-07	30
8,1	2,8466E-06	6,05605E-07	60
4,2	1,3571E-06	1,24316E-07	115
4,1	1,4233E-06	1,25705E-07	120
19,1	0,000003641	1,0862E-06	3
18,5	3,5748E-06	1,07648E-06	4
17,7	3,3762E-06	1,10287E-06	5
17,3	3,2769E-06	8,841E-07	7
16,2	3,5748E-06	7,7923E-07	10
14,5	3,5417E-06	7,71591E-07	15
13,7	3,5748E-06	8,6257E-07	20
13,1	3,6079E-06	8,39652E-07	25
11,8	3,7403E-06	9,32715E-07	30
7,6	3,4424E-06	5,8477E-07	60
5,3	2,1184E-06	2,78495E-07	115
5,3	1,8205E-06	2,30574E-07	120
15	2,3832E-06	0,00000087	3
15,5	2,3832E-06	8,75071E-07	4
15,2	2,4825E-06	7,81314E-07	5
15,2	2,6811E-06	7,27837E-07	7
14,9	2,2839E-06	6,52831E-07	10
13,8	2,6149E-06	7,27143E-07	15
12,4	2,4163E-06	7,27143E-07	20

11,3	0,000002648	7,06308E-07	25
10	2,5156E-06	5,60462E-07	30
5,9	1,9198E-06	2,59049E-07	60
4,5	1,0923E-06	9,51466E-08	115
4,4	0,000000993	9,16741E-08	120
10	1,7212E-06	3,51418E-07	3
11,5	2,2177E-06	5,84075E-07	4
11,3	2,2177E-06	6,89639E-07	5
11,2	2,2839E-06	6,66026E-07	7
11	0,000002317	5,30599E-07	10
11,1	2,2177E-06	4,01422E-07	15
10,4	2,0191E-06	3,45861E-07	20
9,6	2,2508E-06	3,49334E-07	25
9	2,1184E-06	2,54882E-07	30
6,9	1,7874E-06	1,45151E-07	60
5	1,2909E-06	9,65356E-08	115
4,9	1,2578E-06	8,40346E-08	120

Bibliography

Adaszewski, Stanisław et al. (May 15, 2018). "Simultaneous estimation of population receptive field and hemodynamic parameters from single point BOLD responses using Metropolis-Hastings sampling". In: *NeuroImage* 172, pp. 175–193. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2018.01.047 (cit. on p. 26).

Agbaje, Olorunsola F et al. (2003). "Bayesian hierarchical approach to estimate insulin sensitivity by minimal model". In: *Clinical science* 105.5, pp. 551–560 (cit. on p. 5).

Aït-Sahalia, Yacine (2002). "Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach". In: *Econometrica* 70.1, pp. 223–262 (cit. on pp. 19, 33, 38).

Aït-Sahalia, Yacine (2002). "[Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes]: Comment". In: *Journal of Business & Economic Statistics* 20.3, pp. 317–321 (cit. on p. 33).

Ait-Sahalia, Yacine et al. (2008). "Closed-form likelihood expansions for multivariate diffusions". In: *The Annals of Statistics* 36.2, pp. 906–937 (cit. on pp. 33, 38, 45, 46).

Alberti, Kurt George Matthew Mayer and PZ ft Zimmet (1998). "Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation". In: *Diabetic medicine* 15.7, pp. 539–553 (cit. on p. 58).

Allen, Edward (2007a). *Modeling with Itô stochastic differential equations*. Vol. 22. Springer Science & Business Media (cit. on p. 11).

– (2007b). *Modeling with Itô stochastic differential equations*. Vol. 22. Springer Science & Business Media (cit. on p. 19).

Andersen, Kim E and Malene Højbjerre (2005). "A population-based Bayesian approach to the minimal model of glucose and insulin homeostasis". In: *Statistics in medicine* 24.15, pp. 2381–2400 (cit. on pp. 6, 42).

Andersson, Hakan and Tom Britton (2012). *Stochastic epidemic models and their statistical analysis*. Vol. 151. Springer Science & Business Media (cit. on p. 35).

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein (2010). "Particle markov chain monte carlo methods". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72.3, pp. 269–342 (cit. on p. 50).

Arminger, Gerhard and Bengt O. Muthén (Sept. 1, 1998). "A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the metropolis-hastings algorithm". In: *Psychometrika* 63.3, pp. 271–300. ISSN: 1860-0980. DOI: 10.1007/BF02294856 (cit. on p. 26).

Aït-Sahalia, Yacine (Dec. 5, 2009). "Estimating and Testing Continuous-Time Models in Finance: The Role of Transition Densities". In: *Annual Review of Financial Economics* 1.1, pp. 341–359. ISSN: 1941-1367, 1941-1375. DOI: 10.1146/annurev.financial.050808.114424 (cit. on p. 19).

Aït-Sahalia, Yacine and Thomas Robert Hurd (Jan. 1, 2016). "Portfolio Choice in Markets with Contagion". In: *Journal of Financial Econometrics* 14.1, pp. 1–28. ISSN: 1479-8409. DOI: 10.1093/jjfinec/nbv024 (cit. on p. 19).

Aït-Sahalia, Yacine and Jean Jacod (Dec. 2012). "Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data". In: *Journal of Economic Literature* 50.4, pp. 1007–1050. ISSN: 0022-0515. DOI: 10.1257/jel.50.4.1007 (cit. on p. 19).

Aït-Sahalia, Yacine and Dacheng Xiu (Dec. 2017). "Using principal component analysis to estimate a high dimensional factor model with high-frequency data". In: *Journal of Econometrics* 201.2, pp. 384–399. ISSN: 03044076. DOI: 10.1016/j.jeconom.2017.08.015 (cit. on p. 19).

Aït-Sahalia, Yacine et al. (Oct. 2, 2017). "Estimation of the Continuous and Discontinuous Leverage Effects". In: *Journal of the American Statistical Association* 112.520, pp. 1744–1758. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.2016.1240082 (cit. on p. 19).

Bakrim, F and H El Maroufy (2019). "A review on estimation methods of NonLinear Mixed Effects model with Stochastic Differential Equations, application to three-dimensional Ornstein-Uhlenbeck process." In: *Moroccan Journal of Quantitative and Qualitative Research* 1.1, pp. 25–48 (cit. on p. 33).

Becker, Niels G (1977). "On a general stochastic epidemic model". In: *Theoretical Population Biology* 11.1, pp. 23–36 (cit. on p. 35).

Bergman, Richard N, Lawrence S Phillips, and Claudio Cobelli (1981). "Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose." In: *The Journal of clinical investigation* 68.6, pp. 1456–1467 (cit. on pp. 60, 61, 63).

Bergman, Richard N et al. (1979a). "Quantitative estimation of insulin sensitivity." In: *American Journal of Physiology-Endocrinology And Metabolism* 236.6, E667 (cit. on pp. 58, 61).

Bergman, RN et al. (1979b). "Identification of a minimal model of glucose disappearance for estimating insulin sensitivity". In: *IFAC Proceedings Volumes* 12.8, pp. 883–890 (cit. on pp. 60, 61, 63).

Brandt, Michael W and Pedro Santa-Clara (2002). "Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets". In: *Journal of financial economics* 63.2, pp. 161–210 (cit. on pp. 6, 38).

Carter, C. K. and R. Kohn (Sept. 1, 1994). "On Gibbs sampling for state space models". In: *Biometrika* 81.3, pp. 541–553. ISSN: 0006-3444. DOI: 10.1093/biomet/81.3.541 (cit. on p. 26).

Casella, George and Edward I. George (Aug. 1, 1992). "Explaining the Gibbs Sampler". In: *The American Statistician* 46.3, pp. 167–174. ISSN: 0003-1305. DOI: 10.1080/00031305.1992.10475878 (cit. on p. 25).

Cauchemez, S. et al. (2004). "A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data". In: *Statistics in Medicine* 23.22, pp. 3469–3487. ISSN: 1097-0258. DOI: 10.1002/sim.1912 (cit. on p. 26).

Chib, Siddhartha and Edward Greenberg (Feb. 27, 2012). "Understanding the Metropolis-Hastings Algorithm". In: *The American Statistician* (cit. on p. 26).

Choi, Boseung and Grzegorz A Rempala (2011). "Inference for discretely observed stochastic kinetic networks with applications to epidemic modeling". In: *Biostatistics* 13.1, pp. 153–165 (cit. on p. 6).

Cobelli, Claudio et al. (2009). "Diabetes: models, signals, and control". In: *IEEE reviews in biomedical engineering* 2, p. 54 (cit. on pp. 61, 63).

Damlen, P., J. Wakefield, and S. Walker (1999). "Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61.2, pp. 331–344. ISSN: 1467-9868. DOI: 10.1111/1467-9868.00179 (cit. on p. 26).

D'Argenio, David Z and Kyungsoo Park (1997). "Uncertain pharmacokinetic/pharmacodynamic systems: design, estimation and control". In: *Control Engineering Practice* 5.12, pp. 1707–1716 (cit. on p. 11).

Davidian, Marie and David M Giltinan (2003). "Nonlinear models for repeated measurement data: an overview and update". In: *Journal of agricultural, biological, and environmental statistics* 8.4, p. 387 (cit. on p. 39).

De Gaetano, Andrea and Ovide Arino (2000). "Mathematical modelling of the intravenous glucose tolerance test". In: *Journal of mathematical biology* 40.2, pp. 136–168 (cit. on pp. 59, 62, 74).

De Jong, Kenneth A (2006). *Evolutionary computation: a unified approach*. MIT press (cit. on pp. 41, 42).

Delattre, Maud and Marc Lavielle (2013). "Coupling the SAEM algorithm and the extended Kalman filter for maximum likelihood estimation in mixed-effects diffusion models". In: (cit. on pp. 50, 51).

Demiris, Nikolaos and Philip D. O'Neill (2005). "Bayesian inference for stochastic multitype epidemics in structured populations via random graphs". In: *Journal of the*

Royal Statistical Society: Series B (Statistical Methodology) 67.5, pp. 731–745. ISSN: 1467-9868. DOI: 10.1111/j.1467-9868.2005.00524.x (cit. on p. 26).

Dempster, Arthur P, Nan M Laird, and Donald B Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1, pp. 1–22 (cit. on p. 50).

Ditlevsen, Susanne and Andrea De Gaetano (2005a). "Mixed effects in stochastic differential equation models". In: *REVSTAT-Statistical Journal* 3.2, pp. 137–153 (cit. on pp. 33, 70).

– (2005b). "Stochastic vs. deterministic uptake of dodecanedioic acid by isolated rat livers". In: *Bulletin of mathematical biology* 67.3, pp. 547–561 (cit. on p. 6).

– (2005c). "Stochastic vs. deterministic uptake of dodecanedioic acid by isolated rat livers". In: *Bulletin of mathematical biology* 67.3, pp. 547–561 (cit. on pp. 11, 32).

Ditlevsen, Susanne, Kay-Pong Yip, and Niels-Henrik Holstein-Rathlou (2005). "Parameter estimation in a stochastic model of the tubuloglomerular feedback mechanism in a rat nephron". In: *Mathematical biosciences* 194.1, pp. 49–69 (cit. on p. 11).

Ditlevsen, Susanne et al. (2007). "Parameter estimation of feedback gain in a stochastic model of renal hemodynamics: differences between spontaneously hypertensive and Sprague-Dawley rats". In: *American Journal of Physiology-Renal Physiology* 292.2, F607–F616 (cit. on pp. 6, 10, 32).

Donnet, Sophie and Adeline Samson (2008). "Parametric inference for mixed models defined by stochastic differential equations". In: *ESAIM: Probability and Statistics* 12, pp. 196–218 (cit. on pp. 33, 50, 51).

– (2011). "EM algorithm coupled with particle filter for maximum likelihood parameter estimation of stochastic differential mixed-effects models". In: (cit. on pp. 50, 51).

– (2013). "A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models". In: *Advanced drug delivery reviews* 65.7, pp. 929–939 (cit. on pp. 6, 10, 32, 33).

Durham, Garland B and A Ronald Gallant (2002). "Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes". In: *Journal of Business & Economic Statistics* 20.3, pp. 297–338 (cit. on p. 38).

Egorov, Alexei V, Haitao Li, and Yuewu Xu (2003). "Maximum likelihood estimation of time-inhomogeneous diffusions". In: *Journal of Econometrics* 114.1, pp. 107–139 (cit. on pp. 46, 65).

El Maroufy, Hamid, Lahcen Omari, and Ziad Taib (2012). "Transition Probabilities for Generalized SIR Epidemic Model". In: *Stoch. Model.* 28.1, pp. 15–28. ISSN: 1532-6349. DOI: 10.1080/15326349.2011.614201 (cit. on p. 23).

Elerian, Ola et al. (1998). "A note on the existence of a closed form conditional transition density for the Milstein scheme". In: *Economics discussion paper*, W18 (cit. on p. 27).

Eraker, Bjørn (2001). "MCMC analysis of diffusion models with application to finance". In: *Journal of Business & Economic Statistics* 19.2, pp. 177–191 (cit. on pp. 82, 83).

Eraker, Bjørn and Jiakou Wang (Aug. 2015). "A non-linear dynamic model of the variance risk premium". In: *Journal of Econometrics* 187.2, pp. 547–556. ISSN: 03044076. DOI: 10.1016/j.jeconom.2015.02.038 (cit. on p. 19).

Eraker, Bjørn and Yue Wu (July 2017). "Explaining the negative returns to volatility claims: An equilibrium approach". In: *Journal of Financial Economics* 125.1, pp. 72–98. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2017.04.007 (cit. on p. 19).

Fadwa, Bakrim, Hamid El Maroufy, and Hassan Ait Mousse (2020). "Simulation and Parametric Inference of a Mixed Effects Model with Stochastic Differential Equations Using the Fokker-Planck Equation Solution". In: *Numerical Modeling and Computer Simulation*. IntechOpen (cit. on p. 69).

Favetto, Benjamin and Adeline Samson (2010). "Parameter estimation for a bidimensional partially observed Ornstein–Uhlenbeck process with biological application". In: *Scandinavian Journal of Statistics* 37.2, pp. 200–220 (cit. on p. 70).

Ferrante, L, S Bompadre, and L Leone (2003). "A stochastic compartmental model with long lasting infusion". In: *Biometrical Journal: Journal of Mathematical Methods in Biosciences* 45.2, pp. 182–194 (cit. on p. 11).

Fröberg, Carl-Erik (1985). *Numerical mathematics: theory and computer applications*. Benjamin-Cummings Publishing Co., Inc. (cit. on pp. 33, 39).

Geman, S. and D. Geman (Nov. 1984). "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6.6, pp. 721–741. ISSN: 0162-8828. DOI: 10.1109/TPAMI.1984.4767596 (cit. on p. 26).

Geweke, John and Hisashi Tanizaki (Aug. 28, 2001). "Bayesian estimation of state-space models using the Metropolis–Hastings algorithm within Gibbs sampling". In: *Computational Statistics & Data Analysis* 37.2, pp. 151–170. ISSN: 0167-9473. DOI: 10.1016/S0167-9473(01)00009-3 (cit. on p. 26).

Gilks, W. R., N. G. Best, and K. K. C. Tan (1995). "Adaptive Rejection Metropolis Sampling within Gibbs Sampling". In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 44.4, pp. 455–472. ISSN: 0035-9254. DOI: 10.2307/2986138 (cit. on p. 26).

Golberg, David E (1989). "Genetic algorithms in search, optimization, and machine learning. Addion Wesley". In: *Reading* (cit. on p. 41).

Griewank, Andreas (2000). "Evaluating derivatives: principles and techniques of algorithmic differentiation". In: (cit. on p. 40).

Guedj, Jérémie, Rodolphe Thiébaut, and Daniel Commenges (2007). "Maximum likelihood estimation in dynamical models of HIV". In: *Biometrics* 63.4, pp. 1198–1206 (cit. on pp. 10, 32).

Hastings, W. K. (Apr. 1, 1970). "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1, pp. 97–109. ISSN: 0006-3444. DOI: 10.1093/biomet/57.1.97 (cit. on p. 26).

Hitchcock, David B. (Nov. 1, 2003). "A History of the Metropolis–Hastings Algorithm". In: *The American Statistician* 57.4, pp. 254–257. ISSN: 0003-1305. DOI: 10.1198/0003130032413 (cit. on p. 26).

Hurn, A Stan, Kenneth A Lindsay, and Vance L Martin (2003). "On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations". In: *Journal of Time Series Analysis* 24.1, pp. 45–63 (cit. on pp. 34, 38).

Jeliazkov, Ivan (Mar. 1, 2001). "Marginal Likelihood From the Metropolis–Hastings Output AU - Chib, Siddhartha". In: *Journal of the American Statistical Association* 96.453, pp. 270–281. ISSN: 0162-1459. DOI: 10.1198/016214501750332848 (cit. on p. 26).

Jelliffe, Roger, Alan Schumitzky, and Michael Van Guilder (2000). "Population pharmacokinetics/pharmacodynamics modeling: parametric and nonparametric methods". In: *Therapeutic drug monitoring* 22.3, pp. 354–365 (cit. on p. 32).

Klim, Søren et al. (2009). "Population stochastic modelling (PSM)—an R package for mixed-effects models based on stochastic differential equations". In: *Computer methods and programs in biomedicine* 94.3, pp. 279–289 (cit. on p. 53).

Klöden, Peter E and Eckhard Platen (1992). "Numerical Solution of Stochastic Differential Equations". In: (cit. on pp. 36, 72, 74).

Kouritsin, M.A. and Li Deli (Jan. 2000). "On explicit solutions to stochastic differential equations". In: *Stochastic Analysis and Applications* 18.4, pp. 571–580. ISSN: 0736-2994, 1532-9356. DOI: 10.1080/07362990008809686 (cit. on p. 21).

Krommer, Arnold R and Christoph W Ueberhuber (1998). *Computational integration. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998*. Tech. rep. ISBN 0-89871-374-9 (cit. on pp. 33, 39).

Øksendal, Bernt (2003). *Stochastic Differential Equations: An Introduction with Applications*. 6th ed. Universitext. Berlin Heidelberg: Springer-Verlag. ISBN: 978-3-540-04758-2 (cit. on pp. 15, 18, 19, 27, 35, 83, 91).

Kuhn, Estelle and Marc Lavielle (2005). "Maximum likelihood estimation in nonlinear mixed effects models". In: *Computational Statistics & Data Analysis* 49.4, pp. 1020–1038 (cit. on pp. 10, 32, 49).

Kutoyants, Yury A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. Springer London. 482 pp. ISBN: 1-4471-3866-X (cit. on pp. 23, 93).

Lánskỳ, Petr, Vera Lánská, and Michael Weiss (2004). "A stochastic differential equation model for drug dissolution and its parameters". In: *Journal of controlled release* 100.2, pp. 267–274 (cit. on p. 6).

Lindstrom, Mary J and Douglas M Bates (1990). "Nonlinear mixed effects models for repeated measures data". In: *Biometrics*, pp. 673–687 (cit. on pp. 33, 52).

Liptser, Robert S. and Albert N. Shiryaev (2001). *Statistics of Random Processes*. Vol. 6. Springer Berlin Heidelberg. 427 pp. ISBN: 978-3-642-08366-2. DOI: 10.1007/978-3-662-10028-8 (cit. on pp. 23, 93).

Lo, Andrew W (1988). "Maximum likelihood estimation of generalized Itô processes with discretely sampled data". In: *Econometric Theory* 4.2, pp. 231–247 (cit. on pp. 33, 38, 43, 47).

Makroglou, Athena, Jiaxu Li, and Yang Kuang (2006). "Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview". In: *Applied numerical mathematics* 56.3-4, pp. 559–573 (cit. on p. 74).

Mamontov, E (2008). "Dynamic-equilibrium solutions of ordinary differential equations and their role in applied problems". In: *Applied mathematics letters* 21.4, pp. 320–325 (cit. on p. 9).

Martin, Blaise C et al. (1992). "Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study". In: *The Lancet* 340.8825, pp. 925–929 (cit. on p. 59).

Metropolis, Nicholas et al. (June 1, 1953). "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6, pp. 1087–1092. ISSN: 0021-9606. DOI: 10.1063/1.1699114 (cit. on p. 26).

Michalewicz, Zbigniew (1992). "Genetic algorithms+ Data structures= Evolution program". In: *Artificial Intelligence, Berlin: Springer, 1992* (cit. on p. 41).

Mortensen, Stig B et al. (2007). "A matlab framework for estimation of NLME models using stochastic differential equations". In: *Journal of Pharmacokinetics and Pharmacodynamics* 34.5, pp. 623–642 (cit. on p. 53).

Nicolau, João (2002). "A new technique for simulating the likelihood of stochastic differential equations". In: *The Econometrics Journal* 5.1, pp. 91–103 (cit. on pp. 34, 38).

Oksendal, Bernt (2003). *Stochastic differential equations: an introduction with applications* (cit. on pp. 35, 36).

Overgaard, Rune V et al. (2005). "Non-linear mixed-effects models with stochastic differential equations: implementation of an estimation algorithm". In: *Journal of pharmacokinetics and pharmacodynamics* 32.1, pp. 85–107 (cit. on pp. 32, 33, 49, 52).

Overgaard, Rune Viig et al. (2007). "Pkpd model of interleukin-21 effects on thermoregulation in monkeys—application and evaluation of stochastic differential equations". In: *Pharmaceutical research* 24.2, pp. 298–309 (cit. on p. 6).

Pacini, Giovanni and Richard N Bergman (1986). "MINMOD: a computer program to calculate insulin sensitivity and pancreatic responsivity from the frequently sampled intravenous glucose tolerance test". In: *Computer methods and programs in biomedicine* 23.2, pp. 113–122 (cit. on pp. 58, 59, 61).

Panik, Michael J. (Mar. 15, 2017). *Stochastic Differential Equations: An Introduction with Applications in Population Dynamics Modeling*. 1 edition. Wiley. 294 pp. (cit. on p. 27).

Pedersen, Asger Roer (1995). "A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations". In: *Scandinavian journal of statistics*, pp. 55–71 (cit. on p. 38).

Petrovski, Andrei, Alex Wilson, and John McCall (1998). "Statistical analysis of genetic algorithms and inference about optimal factors". In: *School of Mathematics in Computer Science, Faculty of Science and Technology, The Robert Gordon University, Aberdeen, UK* 1, pp. 1–20 (cit. on p. 41).

Picchini, Umberto and Susanne Ditlevsen (2011). "Practical estimation of high dimensional stochastic differential mixed-effects models". In: *Computational Statistics & Data Analysis* 55.3, pp. 1426–1444 (cit. on pp. 33, 39, 40).

Picchini, Umberto, Susanne Ditlevsen, and Andrea De Gaetano (2006a). "Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations". In: *Journal of mathematical biology* 53.5, pp. 771–796 (cit. on pp. 6, 10, 32).

– (2006b). "Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations". In: *Journal of mathematical biology* 53.5, pp. 771–796 (cit. on p. 11).

– (2008). "Maximum likelihood estimation of a time-inhomogeneous stochastic differential model of glucose dynamics". In: *Mathematical medicine and biology: a journal of the IMA* 25.2, pp. 141–155 (cit. on pp. 65, 70).

Picchini, Umberto, ANDREA DE GAETANO, and Susanne Ditlevsen (2010). "Stochastic differential mixed-effects models". In: *Scandinavian Journal of statistics* 37.1, pp. 67–90 (cit. on pp. 6, 10, 32, 33, 39).

Pillonetto, Gianluigi et al. (2002). "Minimal model SI= 0 problem in NIDDM subjects: nonzero Bayesian estimates with credible confidence intervals". In: *American Journal of Physiology-Endocrinology and Metabolism* 282.3, E564–E573 (cit. on p. 59).

Pinheiro, José and Douglas Bates (2006). *Mixed-effects models in S and S-PLUS*. Springer Science & Business Media (cit. on p. 8).

Pinheiro, José C and Douglas M Bates (1995). "Approximations to the log-likelihood function in the nonlinear mixed-effects model". In: *Journal of computational and Graphical Statistics* 4.1, pp. 12–35 (cit. on p. 39).

Pinheiro, José C and Edward C Chao (2006). "Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models". In: *Journal of Computational and Graphical Statistics* 15.1, pp. 58–81 (cit. on pp. 33, 39, 40).

Porteous, Ian et al. (2008). "Fast Collapsed Gibbs Sampling for Latent Dirichlet Allocation". In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '08. New York, NY, USA: ACM, pp. 569–577. ISBN: 978-1-60558-193-4. DOI: 10.1145/1401890.1401960 (cit. on p. 26).

Pratola, Matthew T. (Sept. 2016). "Efficient Metropolis–Hastings Proposal Mechanisms for Bayesian Regression Tree Models". In: *Bayesian Analysis* 11.3, pp. 885–911. ISSN: 1936-0975, 1931-6690. DOI: 10.1214/16-BA999 (cit. on p. 26).

Ramanathan, Murali (1999a). "A method for estimating pharmacokinetic risks of concentration-dependent drug interactions from preclinical data". In: *Drug Metabolism and disposition* 27.12, pp. 1479–1487 (cit. on p. 11).
– (1999b). "An application of Ito's lemma in population pharmacokinetics and pharmacodynamics". In: *Pharmaceutical research* 16.4, p. 584 (cit. on p. 11).
– (1999c). "An application of Ito's lemma in population pharmacokinetics and pharmacodynamics". In: *Pharmaceutical research* 16.4, pp. 584–586 (cit. on p. 35).
Ripley, Brian D (2009). *Stochastic simulation*. Vol. 316. John Wiley & Sons (cit. on pp. 34, 38).
Risken, Hannes (1996). "Fokker-planck equation". In: *The Fokker-Planck Equation*. Springer, pp. 63–95 (cit. on pp. 42, 43, 54, 86, 87).
Robert, Christian and George Casella (2004). *Monte Carlo Statistical Methods*. 2nd ed. Springer Texts in Statistics. New York: Springer-Verlag. ISBN: 978-0-387-21239-5 (cit. on pp. 26, 27, 49).
Roberts, G. O. and O. Stramer (Oct. 1, 2001). "On inference for partially observed nonlinear diffusion models using the Metropolis–Hastings algorithm". In: *Biometrika* 88.3, pp. 603–621. ISSN: 0006-3444. DOI: 10.1093/biomet/88.3.603 (cit. on p. 26).
S Donnet JL Foulley, A Samson (2010). "Bayesian analysis of growth curves using mixed models defined by stochastic differential equations". In: *Biometrics* 37.2, pp. 200–220 (cit. on p. 49).
Searle, Shayle Robert and Charles E McCulloch (2001a). *Generalized, linear and mixed models*. Wiley (cit. on p. 8).
– (2001b). *Generalized, linear, and mixed models*. Wiley (cit. on pp. 10, 32, 33, 39).
Sheiner, Lewis B and Stuart L Beal (1980). "Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data". In: *Journal of pharmacokinetics and biopharmaceutics* 8.6, pp. 553–571 (cit. on pp. 6, 10, 32).
– (1981). "Evaluation of methods for estimating population pharmacokinetic parameters II. Biexponential model and experimental pharmacokinetic data". In: *Journal of pharmacokinetics and biopharmaceutics* 9.5, pp. 635–651 (cit. on pp. 6, 10, 32).
Shun, Zhenming and Peter McCullagh (1995). "Laplace approximation of high dimensional integrals". In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 749–760 (cit. on p. 40).
Sivanandam, SN and SN Deepa (2008). "Genetic algorithm optimization problems". In: *Introduction to Genetic Algorithms*. Springer, pp. 165–209 (cit. on p. 41).
Sørensen, Helle (2004). "Parametric inference for diffusion processes observed at discrete points in time: a survey". In: *International Statistical Review* 72.3, pp. 337–354 (cit. on p. 11).
Toffolo, Gianna, Fabio De Grandi, and Claudio Cobelli (1995). "Estimation of β-cell sensitivity from intravenous glucose tolerance test C-peptide data: knowledge of the

kinetics avoids errors in modeling the secretion". In: *Diabetes* 44.7, pp. 845–854 (cit. on pp. 60, 61, 63).

Tornøe, Christoffer W et al. (2005). "Stochastic differential equations in NONMEM®: implementation, application, and comparison with ordinary differential equations". In: *Pharmaceutical research* 22.8, pp. 1247–1258 (cit. on pp. 32, 33, 49, 52, 53).

Tornøe, Christoffer Wenzel, Judith L Jacobsen, and Henrik Madsen (2004). "Grey-box pharmacokinetic/pharmacodynamic modelling of a euglycaemic clamp study". In: *Journal of mathematical biology* 48.6, pp. 591–604 (cit. on p. 11).

Vicini, Paolo and Claudio Cobelli (2001). "The iterative two-stage population approach to IVGTT minimal modeling: improved precision with reduced sampling". In: *American Journal of Physiology-Endocrinology and Metabolism* 280.1, E179–E186 (cit. on p. 5).

Vonesh, Edward and Vernon M Chinchilli (1996). *Linear and nonlinear models for the analysis of repeated measurements*. CRC press (cit. on pp. 10, 32).

Walker, Stephen (1996). "An EM algorithm for nonlinear random effects models". In: *Biometrics*, pp. 934–944 (cit. on p. 49).

Wang, Jing (2007). "EM algorithms for nonlinear mixed effects models". In: *Computational statistics & data analysis* 51.6, pp. 3244–3256 (cit. on pp. 10, 32, 50).

Wilkinson, Darren J. (Apr. 18, 2006). *Stochastic Modelling for Systems Biology*. Chapman and Hall/CRC. ISBN: 978-1-4200-1066-4. DOI: 10.1201/9781420010664 (cit. on pp. 25, 26).

Yalçınkaya, Abdullah, Birdal Şenoğlu, and Ufuk Yolcu (2018). "Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm". In: *Swarm and Evolutionary Computation* 38, pp. 127–138 (cit. on p. 41).