Applications of the stochastic Galerkin method to epidemic models with uncertainty in their parameters

David Harman

BSc (Hons)

School of Environment and Science
Griffith Sciences
Griffith University

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

August 2018
Abstract

Infectious diseases are a serious problem throughout the world and are responsible for a large number of deaths annually. It is estimated that tuberculosis was responsible for 1.3 million deaths in 2016 worldwide. Malaria, a vector-transmitted disease (transmitted to humans by mosquitoes), is responsible for almost half a million deaths annually. There are also many sexually transmitted diseases, such as chlamydia, genital herpes and gonorrhea, as well as the much more serious HIV/AIDS.

In order to help prevent the spread of an infectious disease, we first need to understand how the disease is spreading through the population, as well as how fast it is spreading. To do this, we need to build mathematical models for the disease. These are referred to as epidemic models. These models can also help predict the effectiveness of interventions, such as treatment and vaccines.

One of the most widely used methods for constructing an epidemic model is the use of compartmental models. Each person within the population is assigned to a specific compartment, based upon their current status with regard to the disease. As their status changes, for example, if they contract the disease, they are moved to the appropriate compartment. Using the compartment model, a system of ordinary differential equations can be derived that models the disease. As the system of ordinary differential equations is usually non-linear, numerical solvers often need to be used as an analytic solution is rarely obtainable.

While it is usually a relatively easy process to derive a model for a particular disease, the parameters within these compartment models are rarely known and usually have to be estimated. Even for well-studied or seasonal diseases, these parameter values are usually not known with certainty and are instead given as probability distributions or simply as a plausible range of values. Since the parameter values are not known with certainty, it is important for this uncertainty to be included in the model. Not including the uncertainty in the model could lead to inaccurate and misleading predictions.

This thesis looks at using the stochastic Galerkin method to incorporate uncertainty in epidemic modelling. While there is extensive literature on the stochastic Galerkin method, there has been very little research into its applications to epidemic modelling.

The stochastic Galerkin method employs a spectral approach, using orthogo-
nal polynomials as basis functions. Substituting the spectral expansion into the epidemic model, and performing a Galerkin projection, results in a deterministic system of differential equations. While this system of differential equations is usually much larger than the original system of differential equations, it is deterministic and hence needs only to be solved once. This results in a significant speed increase over sampling techniques. From the spectral expansion, the mean and variance can quickly be calculated without the need for sampling.

To demonstrate how to apply the stochastic Galerkin method to epidemic models, as well as to analyse its accuracy, in this thesis the stochastic Galerkin method is applied to the SIR epidemic model using different combinations of uncertain parameters and initial conditions. Uniform, gamma and normal distributions are considered for the uncertain parameters. The stochastic Galerkin method is shown to produce accurate results when compared to those obtained using Monte Carlo sampling, while also obtaining the results much more quickly.

While applying the stochastic Galerkin method to various epidemic models, it was found that the stochastic Galerkin method could not always be calculated and could instead ‘blow-out’. While this has been noted as a possibility by some researchers when using unrealistic parameter values, it is shown here that these blow-outs can occur even when care is taken to ensure that the parameter values are sensible and realistic.

These blow-outs are a significant draw-back to the stochastic Galerkin method and have not previously been investigated. By using low order stochastic Galerkin expansions, this work has found that the stochastic Galerkin solution cannot be calculated if the stable attractor is not present in the resulting system of deterministic differential equations. It is shown that for low order expansions, it is possible to analytically predict a range of parameter values for which the stochastic Galerkin solution can be calculated, but unfortunately this does not help predict where blow-outs would occur for higher order expansions.

A data set from an epidemic that went through a small boarding school was then investigated. Rather than simply considering the parameter values that produced the ‘best’ fit to the data, a range of parameter values that resulted in a reasonably accurate fit to the data was instead considered. It was found that this range of plausible values formed a simple closed shape on a 2D plot. It was then shown that by simply finding the border of this shape, probability distributions of the uncertain parameters could be calculated. This eliminated the need to test many of the parameter values.

As these probability distributions were non-standard, this study next extended the stochastic Galerkin method to work with probability distributions other than uniform, gamma, normal and beta distributions. The orthogonal polynomials associated with the non-standard probability distributions could be quickly calcu-
lated using the Gram-Schmidt orthonormalisation method. Extending the stochastic Galerkin method to work with non-standard probability distributions allowed for much greater flexibility in representing the uncertainty in the parameters. Next, using the stochastic Galerkin method, predictions were calculated about what might have happened if the disease had spread outside of the boarding school.

The investigation of the boarding school epidemic was then extended by assuming that the uncertain parameters were no longer independent. While it can be argued that the parameters in an epidemic model are independent, based upon the real world conditions that they are attempting to model, it was clear from the ranges of plausible values that the parameters were dependent, and should be treated that way. Dependent distributions for the parameters were determined and the associated orthogonal polynomials derived. The stochastic Galerkin method was applied and predictions obtained. While the predictions were not particularly accurate, it is hoped that this approach will be helpful and could be studied further to increase the accuracy of its predictions.

Finally, predictions were made on each day of the boarding school epidemic using only the data that would have been available on that day. These predictions were then compared to known data points. In order to obtain the predictions, probability distributions for the uncertain parameters needed to be determined on each day of the epidemic and the associated orthogonal polynomials derived before the stochastic Galerkin method could be applied. It was found that predictions made before the peak of the epidemic had very large variances, but predictions made after the peak of the epidemic were relatively accurate and had much smaller variances.
Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Signed: ____________________________

David Harman
Acknowledgments

Throughout my undergraduate degree, Honours and PhD candidature, there have been a number of people who have selflessly assisted me during my studies. While my true thanks to these people could never be expressed properly in words, I would like to take this opportunity to thank these people for their contribution to my higher educational studies and recognise that I would not be here if it was not for them.

First and foremost I would like to thank my supervisor, Associate Professor Peter Johnston, for his unwavering support and guidance. As an undergraduate student, I always enjoyed Peter’s courses. He was also my supervisor during my Honours year. Peter has always been generous with his time and assistance. He always made time for us to meet and discuss our research. Without his support, this thesis would not have been possible.

I would also like to thank Dr Barbara Johnston for her help and advice, not only with my thesis, but also with my confirmation.

I am very grateful to Griffith University for their support during my undergraduate and postgraduate studies as well as the for the financial support provided during my candidature.

This research was also supported through an Australian Government Research Training Program Scholarship, for which I am very grateful.

My heartfelt thanks go to my parents, Eric and Sandra, for all of their love and support. I would also like to thank my brothers, Joel and Andrew, for all of their ongoing love and encouragement.
Table of Contents

Abstract i
Statement of Originality iv
Acknowledgments v
Table of Contents vi
List of Figures ix
List of Tables xix
Publications arising from this thesis xx

1 Introduction 1

2 Literature review and background 5
  2.1 Deterministic compartmental epidemic models 5
      2.1.1 SI model 6
      2.1.2 SIS model 9
      2.1.3 SIR model 12
      2.1.4 SIRS model 16
      2.1.5 SEIR model 19
      2.1.6 Including births and deaths in an SIR model 21
      2.1.7 Further extensions to deterministic compartment models 23
  2.2 Background on stochastic Galerkin method 25
      2.2.1 Wiener-Hermite expansion 25
      2.2.2 Generalised polynomial chaos 27
      2.2.3 Stochastic Galerkin method 29

3 Applying the stochastic Galerkin method to the SIR model 32
  3.1 Representing uncertainty in parameters and initial conditions 33
  3.2 SIR model with uncertainty in the transmission rate ($\beta$) 34
  3.3 Using the stochastic Galerkin method 37
### Table of Contents

3.3.1 Determining mean and variance from the stochastic Galerkin solution ................................................. 42  
3.4 Discussion of Roberts’ paper .......................................................... 46  
3.5 SIR model with uncertainty in the initial conditions ......................... 50  
3.6 SIR model with uncertainty in two parameters ............................... 52  
3.7 Determination of Sobol indices .................................................... 59  
3.8 Comparison of results ................................................................ 60  
3.8.1 Comparison of results: uniform distributions ......................... 61  
3.8.2 Comparison of results: gamma distribution .............................. 62  
3.8.3 Comparison of results: normal distribution ............................... 66  
3.8.4 Comparison of computation time vs accuracy ......................... 73  
3.9 Discussion ................................................................................. 74  

4 Issues with ‘blow-outs’  ................................................................... 77  
4.1 Example of a blow-out .................................................................. 77  
4.2 Investigation of blow-outs ............................................................. 83  
4.3 Blow-outs in the SIS model .......................................................... 84  
4.3.1 Gamma distribution ................................................................ 84  
4.3.2 Analytically calculating equilibrium points ............................ 88  
4.3.3 Uniform distribution ............................................................... 92  
4.3.4 Normal distribution ................................................................ 99  
4.4 Returning to SIR model ............................................................... 103  
4.5 Discussion ................................................................................. 107  

5 Applying the stochastic Galerkin method to a data set from an epidemic ......................................................... 110  
5.1 Boarding house influenza epidemic ............................................. 111  
5.2 Finding the ‘best’ values for the parameters β and γ .................. 111  
5.3 Finding plausible values for β and γ ............................................. 115  
5.4 Deriving associated orthogonal polynomials ............................... 118  
5.5 Applying the stochastic Galerkin method .................................... 120  
5.6 Investigating other error thresholds ............................................ 121  
5.7 Determining plausible values for β and γ ................................. 126  
without approximate Bayesian computation ................................. 126  
5.8 Further speed increases ............................................................... 132  
5.9 Discussion ................................................................................. 135  

6 Dependent Distributions ................................................................ 136  
6.1 Determining a dependent probability distribution for β and γ .... 137  
6.2 Applying the stochastic Galerkin method .................................... 142  
6.3 Discussion ................................................................................. 146
List of Figures

2.1 SI compartment model ....................................................... 7
2.2 Example SI Model. \( \beta = 1.5 \). Initially 5% of the population was infected. ......................................................... 8
2.3 SIS compartment model ....................................................... 9
2.4 Example SIS Model. \( \beta = 4, \gamma = 1, R_0 = 4, I_{\text{initial}} = 0.05 \). ............ 11
2.5 Example SIS Model. \( \beta = 0.5, \gamma = 1, R_0 = 0.5, I_{\text{initial}} = 0.4 \). ............ 11
2.6 SIR compartment model ....................................................... 13
2.7 Example SIR Model. \( \beta = 4, \gamma = 1, R_0 = 4, I_{\text{initial}} = 0.05 \). ............ 15
2.8 Example SIR Model. \( \beta = 0.5, \gamma = 1, R_0 = 0.5, I_{\text{initial}} = 0.05 \). ............ 15
2.9 Phase plane for SIR model. \( \beta = 4, \gamma = 1, R_0 = 4 \). Blue lines are sample solutions. For \( S < 1/R_0 \) (dotted line) \( I = 0 \) is a stable equilibrium. ............ 16
2.10 SIRS compartment model ..................................................... 17
2.11 Example SIRS Model. \( \beta = 4, \gamma = 1, \delta = 0.4, R_0 = 4, I_{\text{initial}} = 0.05 \). 18
2.12 Example SIRS Model. \( \beta = 0.5, \gamma = 1, \delta = 0.4, R_0 = 0.5, I_{\text{initial}} = 0.05 \). 19
2.13 SEIR compartment model ................................................... 20
2.14 Example SEIR Model. \( \beta = 4, \alpha = 1, \gamma = 1, I_{\text{initial}} = 0.05 \). ............ 21
2.15 SIR compartment model with births and deaths ..................... 22
2.16 Example SIR Model with births and deaths included. \( \beta = 3, \gamma = 1, b = 0.1, I_{\text{initial}} = 0.05 \). 23

3.1 Solutions for SIR model with uncertainty in \( \beta \). Blue are susceptible and red are infected. Solid lines calculated using Monte Carlo sampling \( (\alpha = 1) \) while dashed lines use average value of \( \beta \) \( (\alpha = 0) \). \( \beta_0 = 4, \gamma = 1, \xi \) is a continuous random variable with a uniform distribution on \([-1, 1]\), \( I_{\text{initial}} = 0.01 \). .................. 35
3.2 Probability distributions for \( \beta \) for Figures 3.1 and 3.3. \( \beta_0 = 4, \xi \) is a continuous random variable with a uniform distribution on \([-1, 1]\). .... 36
3.3 Solutions for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. Solid lines calculated using Monte Carlo sampling ($a = 2$) while dashed lines use average value of $\beta$ ($a = 0$). $\beta_0 = 4$, $\gamma = 1$, $\xi$ is a continuous random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. .......................... 38

3.4 Solutions of $S_i(t)$ and $I_i(t)$ for an SIR model with uncertainty in $\beta$. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 2$, $\xi$ is a random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. .......................... 43

3.5 Comparison of stochastic Galerkin solution with Monte Carlo sampling for SIR model with uncertainty in $\beta$. Blue is susceptible and red is infected. Solid lines are Monte Carlo solutions and dashed lines are stochastic Galerkin solutions. $\beta_0 = 4$, $\gamma = 1$, $a = 2$, $\xi$ is a random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. .......................... 45

3.6 Probability distributions for $\beta$, $\gamma$ and $R_0$ with $\beta_0 = 6$, $\gamma_0 = 3$, $a = 3$, $b = 1$. $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. .......................... 49

3.7 Comparison of stochastic Galerkin solution with Monte Carlo sampling for SIR model with uncertainty in the initial conditions. Blue is susceptible and red is infected. Solid lines are Monte Carlo solutions and dashed lines are stochastic Galerkin solutions. $\beta = 4$, $\gamma = 1$, $\xi$ is a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.1 + 0.08 \xi$. .......................... 53

3.8 Probability distributions for $S_{\text{initial}}$ and $I_{\text{initial}}$ for Figure 3.7. .......................... 54

3.9 Probability distributions for $\beta$ and $\gamma$ for Figure 3.10. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$, $b = 0.5$. $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. .......................... 62

3.10 Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$, $b = 0.5$. $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{\text{initial}} = 0.01$. .......................... 63

3.11 Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$, $b = 0.5$. $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{\text{initial}} = 0.01$, $P = 4$. .......................... 64

3.12 Probability distributions for $\beta$ and $\gamma$ for Figure 3.13. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$. $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one. .......................... 66
3.13 Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$, $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one. $I_{\text{initial}} = 0.01$. 

3.14 Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$, $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one. $I_{\text{initial}} = 0.01$, $P = 3$. 

3.15 Probability distributions for $\beta$ and $\gamma$ for Figure 3.16. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$. $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one, respectively. 

3.16 Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$, $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one respectively. $I_{\text{initial}} = 0.01$, $P = 4$. 

3.17 Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$, $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one respectively. $I_{\text{initial}} = 0.01$, $P = 4$. 

3.18 Comparison of computational time against error of the stochastic Galerkin and Monte Carlo solutions for SIR model with uncertainty in $\beta$ and $\gamma$. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi_1$ and $\xi_2$ have either uniform distributions or gamma distributions. $I_{\text{initial}} = 0.01$. 

4.1 Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 0.24$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.01$. 

4.2 Blow-out in the stochastic Galerkin method at approximately $t = 2$. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 0.25$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.01$. 

4.3 Blow-out in the stochastic Galerkin method at approximately $t = 2$. Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 0.25$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.01$. ................. 82

4.4 Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIS model with uncertainty in $\beta$ and $\gamma$. Red are infected. $P = 1$, $\beta_0 = 4$, $\gamma_0 = 1$, $a = b = 1$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.1$. ................. 85

4.5 Blow-out of the stochastic Galerkin solution for SIS model with uncertainty in $\beta$ and $\gamma$. Blue is $I_0(t)$ and red is $I_1(t)$. $P = 1$, $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$, $b = 1.3$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.1$, ....... 86

4.6 Phase planes for SIS model with uncertainty in $\beta$ and $\gamma$ with $P = 1$. Black lines are nullclines ($dI_0/dt = dI_1/dt = 0$). Dark green dashed lines are stable/unstable manifolds. Blue lines are sample solutions. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. ................. 87

4.7 Discriminant ($D$) for SIS model with uncertainty in $\beta$ and $\gamma$. Black line is $D = 0$ (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method numerically blew-out. The stochastic Galerkin solution could be calculated at blue diamond when $a = b = 1$, whereas the red diamond is $a = 1$, $b = 1.3$, where the stochastic Galerkin method blew-out. Bottom graph has larger bounds on $a$ and $b$ compared to top graph. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi$ has a gamma distribution, with scale factor and shape parameter equal to one. ................. 93

4.8 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIS model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. ................. 94
4.9 Discriminant \((D)\) for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\). Black line is \(D = 0\) (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘\times’ shows where the stochastic Galerkin method numerically blew-out. \(\beta_0 = 4, \gamma_0 = 1, \xi\) is a uniformly distributed random variable on \([-1, 1]\). \(I_{\text{initial}} = 0.01\) 

4.10 Phase planes for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\) with \(P = 1\). Middle and bottom phase plane shows enlarged sections of top phase plane. Black lines are nullclines \((dI_0/dt = dI_1/dt = 0)\). Dark green lines are stable/unstable manifolds. Blue lines are sample solutions. \(\beta_0 = 4, \gamma_0 = 1, a = 1, b = 3.2\). \(\xi\) is a uniformly distributed random variable on \([-1, 1]\). 

4.11 Discriminant \((D)\) for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\). Black line is \(D = 0\) (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘\times’ shows where the stochastic Galerkin method numerically blew-out. \(\beta_0 = 4, \gamma_0 = 1, \xi\) has a uniform distribution on \([-1, 1]\). \(I_{\text{initial}} = 0.20\)

4.12 Phase plane for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\) with \(P = 1\). Black lines are nullclines \((dI_0/dt = dI_1/dt = 0)\). Dark green dashed lines are stable/unstable manifolds. Blue lines are sample solutions. \(\beta_0 = 4, \gamma_0 = 1, a = 6, b = 0\). \(\xi\) has a uniform distribution on \([-1, 1]\).

4.13 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\). Blue ‘+’ shows where the stochastic Galerkin solution could be calculated, while red ‘\times’ shows where the stochastic Galerkin method blew-out. \(\beta = 4, \gamma = 1, \xi\) is a uniformly distributed random variable on \([-1, 1]\). 

4.14 Discriminant \((D)\) for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\). Black line is \(D = 0\) (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘\times’ shows where the stochastic Galerkin method numerically blew-out. \(\beta_0 = 4, \gamma_0 = 1, \xi\) is a normally distributed random variable with zero mean and unit variance. \(I_{\text{initial}} = 0.01\)

4.15 Phase planes for \textit{SIS} model with uncertainty in \(\beta\) and \(\gamma\) with \(P = 1\). Black lines are nullclines \((dI_0/dt = dI_1/dt = 0)\). Dark green lines are stable/unstable manifolds. Blue lines are sample solutions. \(\beta_0 = 4, \gamma_0 = 1, a = 0.3\). \(\xi\) is a normally distributed random variable with zero mean and unit variance.
4.16 Discriminant ($D$) for SIS model with uncertainty in $\beta$ and $\gamma$. Black line is $D = 0$ (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method numerically blew-out. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi$ has a normal distribution with zero mean and unit variance. $I_{\text{initial}} = 0.2$  

4.17 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIS model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi$ is a normally distributed random variable with zero mean and unit variance.  

4.18 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIR model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi_1$ and $\xi_2$ are gamma distributed random variables with unit shape and scale parameters. $I_{\text{initial}} = 0.01$.  

4.19 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIR model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{\text{initial}} = 0.01$.  

4.20 Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIR model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi_1$ and $\xi_2$ are normally distributed random variables with zero mean and unit variance. $I_{\text{initial}} = 0.01$.  

5.1 Influenza spreading within a small boarding school. Blue squares are known data points. Green line is the ‘best fit’ using an SIR model with $\beta = 1.665$ and $\gamma = 0.453$.  

5.2 Influenza spreading within a small boarding school. Blue squares are known data points. Green line is the ‘best fit’ using an SIR model with $\beta = 1.665$ and $\gamma = 0.453$. Red line is the ‘best fit’ using an SEIR model with $\beta = 1.804$, $\alpha = 15.487$ and $\gamma = 0.462$.  

5.3 Histograms of plausible $\beta$ and $\gamma$ values ($E_\gamma^\beta < 0.25$) obtained for the boarding house epidemic using approximate Bayesian computation.
5.4 Approximating the probability distributions for $\beta$ and $\gamma$ with polynomials. Blue is a third order approximation, magenta is a fourth order approximation, black is a fifth order approximation, cyan is a sixth order approximation and red is a seventh order approximation.

5.5 Stochastic Galerkin solutions to SIR model using plausible $\beta$ and $\gamma$ values where $E_{\gamma}^\beta < 0.25$. Blue squares are known data points. Green is the ‘best’ fit. Blue uses third order polynomial approximation, magenta uses fourth order, black uses fifth order, cyan uses sixth order and red uses seventh order. Solid lines are the mean solutions and dashed lines are $\pm$ one standard deviation from the mean. (All stochastic Galerkin solutions are extremely similar with red plotted last).

5.6 Histograms of plausible $\beta$ and $\gamma$ values obtained using approximate Bayesian computation for the three different error thresholds.

5.7 Distributions for $R_0$ for the three different error thresholds.

5.8 Approximating the probability distributions for $\beta$ and $\gamma$ using fifth order polynomials. Red is $E_{\gamma}^\gamma < 0.15$, blue is $E_{\gamma}^\gamma < 0.25$ and black is $E_{\gamma}^\gamma < 0.35$.

5.9 Stochastic Galerkin solutions ($P = 3$) to SIR model using plausible $\beta$ and $\gamma$ values. Red is $E_{\gamma}^\beta < 0.15$, blue is $E_{\gamma}^\beta < 0.25$ and black is $E_{\gamma}^\beta < 0.35$. Blue squares are known data points. Green is the ‘best’ fit. Solid lines are the mean solutions and dashed lines are $\pm$ one standard deviation from the mean.

5.10 Heat map of error associated with different $\beta$ and $\gamma$ values for an SIR model compared to known data points.

5.11 Plot of plausible values for $\beta$ and $\gamma$ using $E_{\gamma}^\beta < 0.25$ using a coarse grid. Blue ‘+’ are less than the error threshold while red ‘x’ are higher than the error threshold.

5.12 Histograms of plausible $\beta$ and $\gamma$ values using $E_{\gamma}^\beta < 0.25$ on a course grid.

5.13 Plots of plausible values for $\beta$ and $\gamma$ using different error thresholds. Blue ‘+’ are less than the error threshold while red ‘x’ are higher than the error threshold.

5.14 Approximating the probability distributions for $\beta$ and $\gamma$ using fifth order polynomials. Red is $E_{\gamma}^\gamma < 0.15$, blue is $E_{\gamma}^\gamma < 0.25$ and black is $E_{\gamma}^\gamma < 0.35$. 
5.15 Stochastic Galerkin solutions to SIR model using plausible \( \beta \) and \( \gamma \) values. Red is \( E_\beta^p < 0.15 \), blue is \( E_\gamma^p < 0.25 \) and black is \( E_\gamma^p < 0.35 \). Blue squares are known data points. Green is the ‘best’ fit. Solid lines are the mean solutions and dashed lines are \( \pm \) one standard deviation from the mean.

5.16 Border of plausible values for \( \beta \) and \( \gamma \) at three different error thresholds. Red x have an error higher than the threshold while blue + have an error less than the threshold.

6.1 Heat map of the error (Equation (5.1)) associated with different \( \beta \) and \( \gamma \) values for an SIR model compared to known data points from a boarding school epidemic. Plausible range of \( \gamma \) values when \( \beta = 1.4 \) using a 0.35 error threshold is shown with dashed black lines.

6.2 Error plot (Equation (5.1)) of SIR model for boarding house epidemic for different \( \beta \) and \( \gamma \) values.

6.3 Plot of \( 1/E_\beta^p \) for boarding house epidemic using SIR model.

6.4 Analytic approximation to \( 1/E_\gamma^p \) plot (Figure 6.3) using Equation (6.3).

6.5 Dependent probability distribution, \( w(\xi_1, \xi_2) \) for \( \beta \) and \( \gamma \) were \( \xi_1 \) and \( \xi_2 \) are the random variables associated with \( \beta \) and \( \gamma \) respectively.

6.6 Stochastic Galerkin solutions to SIR model when \( \beta \) and \( \gamma \) have a dependent probability distribution. Blue squares are known data points. Green is the ‘best’ fit. Solid lines are the mean solution and dashed lines are \( \pm \) one standard deviation from the mean. Red used a narrowed \( \beta \) range, while blue used a narrowed \( \gamma \) range.

7.1 Best fit prediction on day 4 of the epidemic. Best fit has parameter values of \( \beta \approx 1.0924 \) and \( \gamma = 0 \) with an error of approximately 0.0025. Blue squares are known data points and blue circles are future data points.

7.2 Error heat map showing ranges of plausible \( \beta \) and \( \gamma \) values on day 4 of the epidemic. Inner contour has an error threshold of 0.0050 and the outer contour has an error of 0.0075. The two contours are very narrow and very close together.

7.3 Plausible \( \beta \) and \( \gamma \) probability distributions on day 4 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively.
7.4 Predictions made on day 4 of the epidemic. Squares are known data points, while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 154
7.5 Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 5 of the epidemic. ................................................................. 155
7.6 Plausible $\beta$ and $\gamma$ probability distributions on day 5 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively. ................................................................. 155
7.7 Predictions made on day 5 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 156
7.8 Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 6 of the epidemic. ................................................................. 157
7.9 Plausible $\beta$ and $\gamma$ probability distributions on day 6 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively. ................................................................. 157
7.10 Predictions made on day 6 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 158
7.11 Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 7 of the epidemic. ................................................................. 159
7.12 Plausible $\beta$ and $\gamma$ probability distributions on day 7 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively. ................................................................. 159
7.13 Predictions made on day 7 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 160
7.14 Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 9 of the epidemic. ................................................................. 160

7.15 Plausible $\beta$ and $\gamma$ probability distributions on day 9 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively. ................................................................. 161

7.16 Predictions made on day 9 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 162

7.17 Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 11 of the epidemic. ................................................................. 162

7.18 Plausible $\beta$ and $\gamma$ probability distributions on day 11 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively. ................................................................. 163

7.19 Predictions made on day 11 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean. ................................................................. 163


**List of Tables**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>A few common distributions and their associated orthogonal polynomials from the Askey scheme. Note that $B$ is the beta function where $B(\alpha + 1, \beta + 1) = \frac{\alpha \beta}{(\alpha + \beta + 1)!}$</td>
<td>29</td>
</tr>
<tr>
<td>3.1</td>
<td>Terms involved in different order expansions when two parameters contain uncertainty in an SIR model</td>
<td>55</td>
</tr>
<tr>
<td>5.1</td>
<td>Number of students infected with influenza at an English boarding school of 763 male students</td>
<td>112</td>
</tr>
</tbody>
</table>
Publications arising from this thesis


D. B. Harman and P. R. Johnston. Using the stochastic Galerkin method as a predictive tool during an epidemic. (*ANZIAM Journal* - Under Revision)

Chapter 1

Introduction

Infectious diseases are a serious problem throughout the world. In 2016, there were an estimated 216 million cases of malaria, which resulted in an estimated 445,000 deaths [1]. Tuberculosis, commonly referred to as TB, is estimated to be responsible for 1.3 million deaths per year [2], while influenza is estimated to be responsible for between 250,000 and 500,000 deaths per year [3]. This makes the study of infectious diseases of great importance.

Depending upon the specific infectious disease, there are different ways it can spread throughout the population. Some diseases, such as influenza, spread simply by coming into close proximity to someone who is already infected. Sexually transmitted diseases, such as chlamydia or gonorrhea, require much closer contact with someone that is infected and the transfer of bodily fluids [4]. Vector transmitted diseases, such as malaria and dengue-fever, are not spread from person to person but are instead transmitted to humans by mosquitoes which have fed on an infected person [5, 6]. As each disease is different, mathematical models can be used to help predict how fast an infectious disease will spread through a population, as well as the effectiveness of potential interventions, such as school closures [7], treatment or vaccines [8, 9, 10].

Compartmental models are one of the most common methods of constructing a mathematical model for a particular disease [11]. In a compartmental model, every stage of the disease is designated its own compartment. Each person within the population is then assigned to a compartment based upon their current status with regard to the disease. For example, a healthy person who has now become infected with the disease is placed into one compartment, while a person who had the disease but has now recovered will be placed in a different compartment. From the compartmental model, a system of differential equations can be derived that models the spread of the disease.

There is extensive literature on modelling epidemics using compartment models. However, the parameters within these models, as well as their initial conditions, are rarely known with certainty and have to be estimated using the data available [12,
As this uncertainty could be the difference between the disease quickly dying out on its own, and a large scale epidemic, it is important for any uncertainty to be incorporated into the model before calculating any predictions.

One of the simplest ways to include this uncertainty is to use Monte Carlo sampling. By choosing appropriate values for the uncertain parameters, based upon their probability distributions, a single realisation of the epidemic model can be calculated. Repeating this process a large number of times allows for the mean prediction to be calculated, as well as its variance. However, due to slow convergence of the Monte Carlo sampling method, this can be computationally expensive \([14, 15, 16]\). More efficient sampling methods such as Latin hypercube sampling \([17]\) or Smolyak sparse grids \([18, 19]\) could instead be used to reduce the computation time. However, these methods still rely on sampling, and therefore the model still needs to be solved numerous times to obtain the mean prediction and its variance. Hence, a non-sampling method would be preferable.

This thesis investigates using the stochastic Galerkin method to incorporate uncertainty into compartmental epidemic models. To incorporate uncertainty, the parameters and initial conditions are expressed as functions of random variables. Therefore the uncertain parameters and initial conditions are now probability distributions rather than constants. The stochastic Galerkin method uses the generalised polynomial chaos expansion to find a spectral solution, utilising orthogonal polynomials from the Askey scheme as basis functions. This leads to a system of deterministic differential equations \([20]\). While uncertainty is represented using random variables, the final system of equations is deterministic and only needs to be solved once. This provides a significant speed increase over sampling methods.

While there is extensive literature on the stochastic Galerkin method \([21, 22, 23, 24, 14, 25]\), there have been very few papers that investigate its applications to epidemic modelling. Roberts \([12]\) applies the stochastic Galerkin method to an \(SIR\) model and looks at incorporating uncertainty into the basic reproduction number. However, the basic reproduction number is a ratio of two parameters within the \(SIR\) model and the influence on the uncertainty by the individual parameters is not considered. Roberts \([26]\) also applies the stochastic Galerkin method to a two-strain influenza model with uncertainty in the cross-immunity of the two strains. Once again, only one parameter within the model is considered to be uncertain.

Hickson and Roberts \([27]\) apply the stochastic Galerkin method to an \(SIR\) model where heterogeneity in the susceptible and infected populations is considered. Rather than assuming that all individuals are equally susceptible to contracting the disease and all infected individuals are equally infectious, an uncertain parameter was used to simulate heterogeneity. As with Roberts’ other papers \([12, 26]\), only a single uncertain parameter was considered. Also, while the stochastic Galerkin method is used, it is simply referred to in the paper as ‘separation of variables’.
Chen-Charpentier and Stanescu [13] apply the stochastic Galerkin method to \textit{SIR} and \textit{SIRS} epidemic models. Two parameters in the \textit{SIR} model and three parameters in the \textit{SIRS} model were considered to contain uncertainty. However, the initial conditions for both models were considered deterministic. Also, the probability distributions for the parameters were simply assumed and not derived from a data set.

Santonja and Chen-Charpentier [28] apply the stochastic Galerkin method to an obesity epidemic model with uncertainty in the transmission parameters. While not an infectious disease model, the obesity model is also derived from a compartmental model and has similar dynamics to an infectious disease epidemic model. However, due to the small number of data points available, the uncertain parameters were simply assumed to be uniformly distributed.

While researchers have explored some applications of the stochastic Galerkin method to epidemic modelling, there are still many applications that have been left unexplored. This thesis looks at extending the applications of the stochastic Galerkin method to epidemic modelling in three main areas:

1. Extending the stochastic Galerkin method to work with uncertain parameters of any continuous probability distribution. The papers listed above only consider uniform, beta, gamma and normal distributions, which are the four distributions associated with generalised polynomial chaos and the stochastic Galerkin method. Additionally, this was done without decomposing the space of random inputs into smaller elements and applying multi-element generalised polynomial chaos [29, 30].

2. Deriving likely probability distributions for the uncertain parameters directly from an epidemic data set and then using the stochastic Galerkin method to obtain predictions. As the parameters are likely to have non-standard distributions, it is important for the stochastic Galerkin method to work with parameters of any probability distribution (as mentioned above).

3. Exploring the ‘blow-outs’ that can occur when using the stochastic Galerkin method. While applying the stochastic Galerkin method to epidemic models, it was found that in some cases the stochastic Galerkin solution would ‘blow-out’ and could not be calculated. As these blow-outs are a significant drawback to the stochastic Galerkin method, it is important to explore the cause of these blow-outs.

The remainder of this thesis is structured as follows:

Chapter 2 provides a background on the most common compartmental epidemic models and discusses their key dynamics. The background of the Weiner-Hermite expansion and generalised polynomial chaos are also presented as they form the
foundation for the stochastic Galerkin method. A simple example of the stochastic Galerkin method is also presented.

In Chapter 3, the stochastic Galerkin method is applied to the classic $SIR$ epidemic model. Uncertainty is considered in multiple parameters as well as the initial conditions. Uniform, gamma and normal probability distributions are considered for the uncertain parameters. Predictions obtained from the stochastic Galerkin solution are compared to those obtained through Monte Carlo sampling.

Chapter 4 starts by showing an example of a blow-out in the stochastic Galerkin method. Blow-outs are then investigated using low order expansions when the uncertain parameters have uniform, gamma and normal probability distributions. The range of blow-outs for low order expansions is calculated analytically and compared to numerical results. Ranges of blow-outs are also determined numerically for higher order stochastic Galerkin expansions.

Chapter 5 investigates a data set from an epidemic that went through a small boarding school in the North of England. Rather than simply considering the parameter values that result in the best fit to the data, a range of parameter values is instead considered that results in a reasonably accurate fit to the data. Probability distributions are then determined for the uncertain parameters. As these probability distributions are likely to be non-standard (not uniform, gamma, beta or normally distributed), the stochastic Galerkin method is extended to use non-standard distributions. The stochastic Galerkin method is then used to calculate predictions of what might have happened if the disease had escaped the boarding school and spread into the surrounding population.

Chapter 6 revisits the boarding school epidemic and now assumes that the uncertain parameters are no longer independent. New, dependent distributions are derived for the uncertain parameters. The stochastic Galerkin method is then applied to obtain updated predictions of what could have occurred in the surrounding population.

In Chapter 7, predictions are made of each day of the boarding school epidemic using only data that would have been available on that day. On each day of the epidemic, probability distributions for each of the uncertain parameters are calculated before applying the stochastic Galerkin method to obtain the current day’s prediction. These predictions are then compared to what actually occurred during the epidemic.

Finally, Chapter 8 provides a summary of the key finding of this thesis including the advantages and disadvantages of applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters and initial conditions. An outline of possible extensions to the work included in this thesis is also presented.
Chapter 2

Literature review and background

The ability to predict the likely course an epidemic will take is extremely important. Because of this, epidemic models have been widely studied. As the dynamics of each disease can vary greatly, there are many different types of compartmental epidemic models. This chapter starts by providing a background on some of the most common models. It is important to understand the dynamics of the deterministic models before uncertainty is included in later chapters.

This chapter also looks at the background of the stochastic Galerkin method. This begins by looking at the Weiner-Hermite expansion as well as generalised polynomial chaos, both of which form the basis for the stochastic Galerkin method.

2.1 Deterministic compartmental epidemic models

One of the most common methods of modelling an epidemic is by using compartmental models. To do this, each person within the population is assigned to a particular compartment based upon their current status with respect to the disease being modelled [31, 32, 12, 28]. Using compartmental models to predict the spread of a disease is credited to Kermack and McKendrick for their paper (published in 1927) A Contribution to the Mathematical Theory of Epidemics [33]. These compartmental models rely on a number of key assumptions:

- the size of the population is large enough that random effects (stochastic effects) of individuals within the population can be ignored,
- each individual within the population is indistinguishable from any other individual,
- the population is well mixed (an individual within the population is able to come into contact with anyone else in the population).
2.1.1 *SI* model

Probably the most simplistic epidemic model that can be constructed using compartments is the *SI* epidemic model. The *SI* epidemic model only uses two compartments to classify individuals within the population [34]. The two compartments are:

- **Susceptible (S):** a *susceptible individual* is not currently infected with the disease, but is capable of becoming infected by coming into contact with someone who is infected,

- **Infected (I):** an *infected individual* is currently infected with the disease and is capable of infecting someone who is susceptible.

In the *SI* model, it is assumed that once an individual has become infected, they do not recover from the disease and will remain infected indefinitely. This is a reasonable assumption assuming that the model is run for a shorter duration than the average recovery time from the disease. As the model is only run for a short period of time, it is common not to include births or deaths in the model [35].

During a given time period, people within the population will come into contact with each other. This may be because they live together, work together or simply come into contact while doing regular activities such as shopping, exercising or commuting.

If an ‘adequate contact’ (two individuals come into sufficiently close proximity for the disease to spread) is made between a susceptible individual and an infected individual, there is a chance, \( p (0 < p \leq 1) \), that the susceptible individual will become infected [32]. At this point, the susceptible individual will move from the susceptible compartment to the infected compartment.

In order to develop a model, assume that each person within the population makes, on average, \( m \) adequate contacts per unit time. Each susceptible individual will therefore make, on average, \( mi/N \) contacts with infected individuals per unit time where \( N \) is the number of individuals in the population and \( i \) is the number of infected individuals within the population. For now, the population size, \( N \), will be considered constant. However, this is not a necessary condition and epidemic models that account for births and deaths are discussed in Section 2.1.6. The number of new individuals becoming infected per unit time is therefore given by \( pmsi/N \), where \( s \) is the number of susceptible individuals in the population.

As \( p \) and \( m \) are rarely known separately, it is common to combine them into a new parameter \( \beta \), where \( \beta = pm \) [26]. As \( p \) and \( m \) are assumed to be positive (and non-zero), it is assumed that \( \beta > 0 \) and is referred to as the ‘transmission coefficient’ [36].

The rate of newly infected individuals, \( di/dt \), is therefore given by \( \beta si/N \). As these newly infected individuals must come from the susceptible compartment, the
differential equations describing the $SI$ model \cite{34} are given by

\[
\frac{ds}{dt} = -\frac{\beta si}{N} \tag{2.1}
\]
\[
\frac{di}{dt} = \frac{\beta si}{N}. \tag{2.2}
\]

Substituting the dimensionless variables $S = s/N$ and $I = i/N$ into Equations (2.1) and (2.2) gives

\[
\frac{dS}{dt} = -\beta SI \tag{2.3}
\]
\[
\frac{dI}{dt} = \beta SI. \tag{2.4}
\]

where $S$ is the fraction of the population which is currently susceptible to the disease and $I$ is the fraction of the population which is currently infected by the disease. Note that $S + I = 1$. A compartment diagram for the $SI$ model can be seen in Figure 2.1.

Figure 2.2 shows an example numeric solution to the $SI$ model (Equations (2.3) and (2.4)) found using the MATLAB solver \texttt{ode45} when 5% of the population was initially infected with the disease. From the figure, it can be seen how the susceptible and infected populations change over time. The way in which the infected population changes over time is often referred to as an ‘epidemic curve’.

Even with only a small fraction of the population initially infected, the entire population becomes infected, leaving no one in the susceptible compartment. This is due to the $SI$ model assuming that an infected person will remain infected and cannot recover from the disease. Also, as deaths are not included in the model, the infected population will continue infecting susceptibles until there are none left.

The $SI$ model (Equations (2.3) and (2.4)) does not need to be solved numerically as an analytic solution can be found. In an $SI$ model, $S + I = 1$ as all individuals must either be part of the susceptible or infected compartments. Therefore,
substituting $S = 1 - I$ into Equation (2.4) gives

$$\frac{dI}{dt} = \beta(1 - I)I = \beta I - \beta I^2. \quad (2.5)$$

By making this substitution, the $SI$ model can be written using a single differential equation. Additionally, it is a differential equation only in $I$. Equation (2.5) is a separable differential equation, commonly known as the logistic equation, and its analytic solution is given by

$$I(t) = \frac{1}{1 + \frac{1}{I_{\text{initial}}} e^{-\beta t}} \quad (2.6)$$

where $I(t = 0) = I_{\text{initial}}$. While it is common to use the notation ‘$I_0$’ for the initial fraction of infected, the notation ‘$I_0$’ will be reserved for use with the stochastic Galerkin method, which will be investigated in the next chapter. Therefore, while not as common, $I_{\text{initial}}$ will be used for the initial fraction of infected.

From Equation (2.6), it can be seen that

$$\lim_{t \to \infty} I(t) = 1$$

assuming $I_{\text{initial}} > 0$. Therefore, if an infected individual were to enter the population, even with a very small value of $\beta$ (assuming $\beta > 0$), the $SI$ model will always result in the entire population eventually becoming infected. While this may seem
Chapter 2. Literature review and background

2.1.2 SIS model

A simple extension to the SI model is the SIS model. In the SIS model, there are still only two compartments (susceptible (S) and infected (I)), but it is assumed that infected individuals will recover from the disease after a finite time. After they have recovered from the disease, they are taken from the infected compartment (I) and returned to the susceptible compartment (S). Once returned to the susceptible compartment, they are immediately capable of being infected again. The compartment diagram for the SIS model can be seen in Figure 2.3.

From the compartment diagram, a system of differential equations can be derived for the SIS model and is given by

\[
\frac{dS}{dt} = -\beta SI + \gamma I \tag{2.7}
\]

\[
\frac{dI}{dt} = \beta SI - \gamma I \tag{2.8}
\]

where \(\beta\) is the transmission coefficient (similar to the SI model) and \(\gamma\) is the ‘recovery rate’. The recovery rate is assumed to be positive (\(\gamma > 0\)) and the quantity \(1/\gamma\) is the average length of time an individual remains infected by the disease [31, 35]. As with the SI model in Section 2.1.1, births and deaths have not been included in this model.

As there are only two compartments in the SIS model, it follows that \(S + I = 1\) as everyone must be either in the susceptible or infected compartments. Substituting
\( S = 1 - I \) into Equation (2.8) gives

\[
\frac{dI}{dt} = \beta(1 - I)I - \gamma I \\
= (\beta - \gamma)I - \beta I^2.
\] (2.9)

Assuming that there is at least one infected individual within the population \( I_{\text{initial}} > 0 \), if \( \beta - \gamma \leq 0 \), then \( dI/dt \) will always be negative. This means the fraction of infected individuals \( I \) is always decreasing. Therefore the relationship between \( \beta \) and \( \gamma \) is very important in determining the long term behaviour of the \textit{SIS} model. If \( \beta/\gamma \leq 1 \) (which is equivalent to \( \beta - \gamma \leq 0 \)), then \( dI/dt < 0 \). This means the fraction of infected individuals \( I \) will continue to decrease until there is no one infected \( (I = 0) \). Alternatively, if \( \beta/\gamma > 1 \), \( dI/dt \) is likely to be positive. The fraction of infected individuals \( I \) will increase which results in an epidemic. The ratio \( \beta/\gamma \) is called the \textit{basic reproduction number} and is given the symbol \( R_0 \) [38, 31]. Therefore:

- if \( R_0 \leq 1 \), the fraction of infected individuals \( I \) will decrease until \( I = 0 \), or
- if \( R_0 > 1 \), there will likely be an epidemic [39].

When a new disease breaks out, a lot of effort is invested into determining \( R_0 \) [12]. Not only does \( R_0 \) indicate whether there is going to be an epidemic, but it also indicates the rate at which it will spread. The larger the value of \( R_0 \), the faster the epidemic will spread through the population.

Figure 2.4 shows an example of an \textit{SIS} model with \( R_0 = 4 \). Initially, there was only a small fraction of the population infected, but this quickly increased before reaching an equilibrium that leaves some of the population uninfected. This is very different from an \textit{SI} model which results in the entire population eventually becoming infected. Diseases such as streptococcus pneumoniae can be modelled with an \textit{SIS} model [40].

Figure 2.5 shows an example of an \textit{SIS} model with \( R_0 = 0.5 \). Even with a significant portion of the population initially infected, the fraction of infected \( (I) \) decreases over time until there are no infected individuals. This is expected as \( R_0 < 1 \).

The equilibrium points of an \textit{SIS} model can be calculated analytically. Starting with Equation (2.9) and letting \( dI/dt = 0 \) gives

\[
\frac{dI}{dt} = 0 = (\beta - \gamma)I - \beta I^2 \\
= [\beta - \gamma - \beta I] I.
\]
Figure 2.4: Example SIS Model. $\beta = 4$, $\gamma = 1$, $R_0 = 4$, $I_{\text{initial}} = 0.05$.

Figure 2.5: Example SIS Model. $\beta = 0.5$, $\gamma = 1$, $R_0 = 0.5$, $I_{\text{initial}} = 0.4$. 

Therefore either \( I = 0 \) (the disease-free equilibrium) or

\[
\beta - \gamma - \beta I = 0.
\]

Rearranging gives

\[
I = \frac{\beta - \gamma}{\beta} = 1 - \frac{1}{R_0}.
\]

(2.10)

Therefore as \( R_0 \) increases, the equilibrium point of the disease in the SIS model has a higher fraction of the population infected. For example, a disease with \( R_0 = 2 \) will reach an equilibrium with half of the population infected, whereas a disease with \( R_0 = 10 \) will not reach an equilibrium until 90% of the population is infected with the disease.

As Equation (2.9) is a separable differential equation, the SIS model can be solved analytically giving

\[
I(t) = \frac{\beta - \gamma}{\beta_{\text{initial}}} e^{-(\beta-\gamma)t} + \beta
\]

(2.11)

where \( I(t = 0) = I_{\text{initial}} \). From this analytic solution, it can be seen that if \( \beta - \gamma \leq 0 \) (ie \( R_0 \leq 1 \)), then

\[
\lim_{t \to \infty} I(t) = 0.
\]

Alternatively, if \( \beta - \gamma > 0 \), which is equivalent to \( R_0 > 1 \), (and assuming \( I_{\text{initial}} > 0 \)) then

\[
\lim_{t \to \infty} I(t) = \frac{\beta - \gamma}{\beta} = 1 - \frac{1}{R_0}
\]

which agrees with Equation (2.10).

### 2.1.3 SIR model

A slightly different extension to the SI model is the SIR model. In the SIR model, infected individuals are able to recover from the disease, similar to the SIS model. However, rather than being placed back into the susceptible compartment, as with the SIS model, recovered individuals are instead assumed to have an immunity to the disease (for at least as long as the duration of the model) and are incapable of becoming infected again. Influenza A, for example, can be modelled with an SIR model [41, 12]. To account for this immunity, a new compartment is needed.

- **Removed (R):** A removed individual can either be someone who
  - has recovered from the disease and either has a permanent immunity to the disease or a temporary immunity that lasts at least as long as the
Chapter 2. Literature review and background

Figure 2.6: SIR compartment model

- is currently infected, but has been isolated so that they are incapable of infecting other susceptibles
- has died from the disease.

A compartment diagram for the SIR model can be seen in Figure 2.6. From the compartment diagram, a system of differential equations can be derived for the SIR model and is given by

\[
dS\frac{dt}{dt} = -\beta SI \quad (2.12)
\]

\[
dI\frac{dt}{dt} = \beta SI - \gamma I \quad (2.13)
\]

\[
dR\frac{dt}{dt} = \gamma I. \quad (2.14)
\]

As with the SIS model, \(\beta\) is the transmission coefficient and \(1/\gamma\) is the average infectious period.

Despite the SIR model’s apparent simplicity, there is no known analytic solution for the SIR model. The system of differential equations for the SIR model can be simplified, as Equation (2.14) is not coupled with Equations (2.12) and (2.13). Therefore, once Equations (2.12) and (2.13) have been solved numerically, \(R\) is simply given by \(1 - S - I\) (as \(S + I + R = 1\)).

To determine an expression for \(R_0\) for the SIR model, it is first assumed that the population is almost entirely susceptible \((S \approx 1)\). For \(dI/dt < 0\), it is required that \(\beta - \gamma < 0\) or \(\beta/\gamma < 1\). Therefore \(R_0 = \beta/\gamma\), which is the same expression for \(R_0\) as the SIS model. If \(R_0 \leq 1\), the fraction of infected individuals \((I)\) will decrease until \(I = 0\). If \(R_0 > 1\), there will most likely be an epidemic. The value of \(R_0\) can also be
thought of as the number of secondary cases due to a single infected person within an entirely susceptible population [42]. Larger values of \( R_0 \) will indicate how fast and how aggressive the epidemic will be. Fast spreading diseases, such as measles, have an estimated \( R_0 \) value between 12 and 18 [9, 43] whereas the value of \( R_0 \) can be as low as 1.5 [44] for influenza.

From Equation (2.14), it can be seen that the only equilibrium point of the SIR model is when \( I = 0 \) (the disease-free equilibrium). At this equilibrium point, \( S + R = 1 \).

The Jacobian for the SIR model is given by

\[
J = \begin{bmatrix}
-\beta I & -\beta S \\
\beta I & \beta S - \gamma
\end{bmatrix}.
\]

The eigenvalues, \( \lambda_1 \) and \( \lambda_2 \), of the disease free equilibrium are then

\[
\begin{align*}
\lambda_1 &= 0 \\
\lambda_2 &= \beta S - \gamma.
\end{align*}
\]

For the disease free equilibrium to be a stable equilibrium, it is required that

\[
\lambda_2 = \beta S - \gamma < 0.
\]

Finally, rearranging gives

\[
S < \frac{\gamma}{\beta} = \frac{1}{R_0}.
\]

Therefore, the disease free equilibrium is not always a stable equilibrium. Initially, the disease free equilibrium might be unstable. However, as the epidemic progresses, susceptibles will become infected, which moves them from the susceptible compartment to the infected compartment. Once the fraction of susceptibles is less than \( 1/R_0 \), the disease will head towards the disease free equilibrium.

An example solution of the SIR model can be seen in Figure 2.7 with \( R_0 = 4 \). As \( R_0 > 1 \), an epidemic occurs and then returns to the disease free equilibrium. Another solution to the SIR model with \( R_0 = 0.5 \) can be seen in Figure 2.8. As \( R_0 < 1 \), there is no epidemic and the fraction of infected individuals continues to decrease until it reaches the disease free equilibrium.

Figure 2.9 shows an example phase plane of an SIR model with \( R_0 = 4 \). Initially, \( I = 0 \) is unstable. However, once \( S < 0.25 \), the disease free equilibrium becomes a stable equilibrium. Note that as \( S + I + R = 1 \), then \( S + I \leq 1 \).

The SIR model can also be used to investigate the effect of vaccinations [9]. Once an individual has been vaccinated against the disease, they are moved directly to the removed group (as they now have immunity to the disease) without needing to go through the infected compartment. If enough of the population can be vaccinated
Figure 2.7: Example SIR Model. $\beta = 4$, $\gamma = 1$, $R_0 = 4$, $I_{initial} = 0.05$.

Figure 2.8: Example SIR Model. $\beta = 0.5$, $\gamma = 1$, $R_0 = 0.5$, $I_{initial} = 0.05$. 
against the disease such that $S_{\text{initial}} < 1/R_0$, the disease free equilibrium will always be a stable equilibrium. Even if $R_0 > 1$, if enough of the population are vaccinated, no epidemic will occur. For high values of $R_0$, this level of vaccination may be hard to achieve. For example, a disease with $R_0 = 10$ will require (over) 90% of the population to be vaccinated. This also assumes that vaccinations are 100% effective against the disease.

2.1.4 SIRS model

A simple extension to the SIR model is the SIRS model. In the SIRS model, it is assumed that the immunity obtained after recovering from the disease is only temporary. Once an individual has lost their temporary immunity, they are returned to the susceptible compartment [45]. A compartment diagram for the SIRS model can be seen in Figure 2.10. From the compartment diagram, a system of differential
The equations for the $SIRS$ model can be derived and is given by

\[
\begin{align*}
\frac{dS}{dt} &= -\beta SI + \delta R \\
\frac{dI}{dt} &= \beta SI - \gamma I \\
\frac{dR}{dt} &= \gamma I - \delta R.
\end{align*}
\]

(2.15) \hspace{2cm} (2.16) \hspace{2cm} (2.17)

The value of $R_0$ is given by $\beta/\gamma$ which is the same as for the $SIS$ and $SIR$ models. While it may seem strange that all three models have the same expression for $R_0$, it is because all three models have the same differential equation for $I$: $dI/dt = \beta SI - \gamma I$. The differences in the models come from the different differential equations for $S$ and $R$.

The equilibrium points of the $SIRS$ model can be found by letting Equations (2.15)-(2.17) equal zero. The equilibrium points are

\[
S = 1, \quad I = 0, \quad R = 0
\]

which is the disease free equilibrium and

\[
S = \frac{\gamma}{\beta}, \quad I = \frac{\delta(1 - \frac{\gamma}{\beta})}{\gamma + \delta}, \quad R = 1 - \frac{\gamma}{\beta} - \frac{\delta(1 - \frac{\gamma}{\beta})}{\gamma + \delta}.
\]

Similar to the $SIS$ model, the disease in the $SIRS$ model may be persistent within the population and reach a stable equilibrium which is not the disease free equilibrium. An example of a disease that can be modelled with the $SIRS$ model is cholera [46].
The Jacobian for the SIRS model is given by

\[
J = \begin{bmatrix}
-\beta I - \delta & -\beta S - \delta \\
\beta I & \beta S - \gamma
\end{bmatrix}.
\]

The eigenvalues, \( \lambda \), of the Jacobian have the characteristic equation

\[
\lambda^2 + \lambda [\beta I - \beta S + \delta + \gamma] + \beta \gamma I - \beta \delta S + \gamma \delta + \beta \delta I = 0.
\]

For the disease free equilibrium \((S = 1, I = 0, R = 0)\), the eigenvalues, \( \lambda_1 \) and \( \lambda_2 \), of the Jacobian are

\[
\lambda_1 = -\delta \quad \lambda_2 = \beta - \gamma.
\]

Therefore, for the disease free equilibrium to be a stable equilibrium, it is required that \( \beta - \gamma < 0 \), which is equivalent to \( R_0 < 1 \).

An example SIRS model is shown in Figure 2.11 with \( R_0 = 4 \). As \( R_0 > 1 \), there is an epidemic. The epidemic reaches its peak and it appears as though the disease is going to return to the disease free equilibrium. However, the number of infected only declines slightly before reaching an equilibrium.

Figure 2.12 shows an example of an SIRS model with \( R_0 = 0.5 \). As \( R_0 < 1 \), there is no epidemic and the disease returns to the disease free equilibrium.
2.1.5 SEIR model

All of the models that have been investigated so far have made the assumption that once an individual has been infected by a disease, they are immediately capable of spreading the disease to susceptible individuals in the population. However, some diseases have a ‘latent time’ after infection. Once an individual has become infected, there is a period of time before they become infectious. To model this behaviour, a new compartment is needed.

- Exposed (E) - An exposed individual has been infected by the disease, but is not yet infectious and cannot transmit the disease to susceptible individuals.

For diseases, such as tuberculosis, which have very long latent periods, it is important to include an exposed compartment in the model [47]. However, diseases with short latent periods, such as chickenpox, can also be modelled with an SEIR model [48].

A compartment diagram for the SEIR model can be seen in Figure 2.13. Individuals move from the exposed compartment to the infected compartment at a rate of $\alpha$ where $1/\alpha$ is the average latent time of the disease. Using the compartment diagram, a system of differential equations can be derived for the SEIR model and
is given by [49]

\[
\frac{dS}{dt} = -\beta SI \\
\frac{dE}{dt} = \beta SI - \alpha E \\
\frac{dI}{dt} = \alpha E - \gamma I \\
\frac{dR}{dt} = \gamma I.
\]  

(2.18) \hspace{1cm} (2.19) \hspace{1cm} (2.20) \hspace{1cm} (2.21)

By letting the left-hand side of Equations (2.18) - (2.21) equal zero, the equilibrium points of the \textit{SEIR} model can be determined. Similar to the \textit{SIR} model, the \textit{SEIR} model only has one equilibrium point, which is the disease free equilibrium \((I = 0, E = 0, S + R = 1)\).

An example \textit{SEIR} model can be seen in Figure 2.14. Initially the fraction of exposed individuals begins to rise while the fraction of susceptible individuals declines. After a short time, the fraction of infected individuals begins to rise as individuals move from the exposed compartment to the infected compartment. The exposed peak occurs shortly before the infected peak. After the epidemic, the disease reaches the disease free equilibrium.
2.1.6 Including births and deaths in an SIR model

All of the compartment epidemic models that have been investigated so far have not included births or (natural) deaths. In many cases, this is a valid assumption as the predictions from the models are for a much shorter time period than the average human lifetime. However, if predictions are needed for a longer period of time, it is necessary to include births and deaths in the model [35].

In this section, the SIR model is extended to include births and deaths. However, any of the compartment epidemic models investigated in this chapter could easily be extended to include births and deaths. For example, a SEIR model could have births and deaths incorporated into the model. Measles is an example of a disease that can be modelled with an SIR model that includes births and deaths [50].

The compartment diagram for the SIR model with births and deaths can be seen in Figure 2.15, where $b$ is the birth rate and $\mu$ is the death rate. It is assumed that all newborns are initially susceptible to the disease and are therefore placed directly into the susceptible compartment. It is also assumed that the the death rate is the same for each of the compartments. This can easily be changed, for example, if infected individuals have a higher mortality rate than those in the other compartments. Note that $s$, $i$ and $r$ are the actual number of individuals in each of the respective compartments and are not the dimensionless variables $S$, $I$ and $R$.

From the compartment diagram, the system of differential equations for the SIR
model with births and deaths is given by

\[
\begin{align*}
\frac{dS}{dt} &= -\frac{\beta si}{N} + bN - \mu s \tag{2.22} \\
\frac{di}{dt} &= \frac{\beta si}{N} - \gamma i - \mu i \tag{2.23} \\
\frac{dr}{dt} &= \gamma i - \mu r \tag{2.24} \\
\frac{dN}{dt} &= (b - \mu)N, \tag{2.25}
\end{align*}
\]

where \(N\) is the total number of individuals in the population. Note that previously \(dN/dt = 0\) as births and deaths were not included so the population size remained constant. The dimensionless variables

\[
S(t) = \frac{s(t)}{N(t)}, \quad I(t) = \frac{i(t)}{N(t)} \quad \text{and} \quad R(t) = \frac{r(t)}{N(t)}
\]

are now reintroduced. Substituting these dimensionless variables into Equations (2.22)-(2.24) gives

\[
\begin{align*}
\frac{dS}{dt} &= -\beta SI + b - bS \\
\frac{dI}{dt} &= \beta SI - \gamma I - bI \\
\frac{dR}{dt} &= \gamma I - bR.
\end{align*}
\]

As \(N(t)\) is now a function of time, rather than a constant as in previous models in this chapter, care needs to be taken when non-dimensionalising the equations.
Chapter 2. Literature review and background

Figure 2.16: Example SIR Model with births and deaths included. $\beta = 3$, $\gamma = 1$, $b = 0.1$, $I_{\text{initial}} = 0.05$.

The above equations are slightly different than what may have been expected from the compartment diagram (Figure 2.15). For example, the death rate, $\mu$, does not appear in the non-dimensionalised equations.

An example SIR model with births and deaths can be seen in Figure 2.16. Initially there is an epidemic as $R_0 > 1$ and nearly all the population is susceptible to the disease. After the epidemic is over, most of the population is now immune to the disease and in the removed compartment. As immune people start to die and newborns join the susceptible compartment, the fraction of the population susceptible to the disease begins to increase. When the fraction of susceptibles reaches a critical point, another epidemic (smaller than the first) begins. The epidemic is much smaller as a large fraction of the population is still immune to the disease from the first epidemic. This is an example of an endemic disease.

2.1.7 Further extensions to deterministic compartment models

Throughout this chapter, some of the simplest and most commonly used deterministic compartment models were investigated. However, these models can easily be extended further, often by adding additional compartments to accommodate the characteristics of the disease being studied. The following are common examples of
extensions to the common compartment models.

- Maternal immunity: This compartment is for newborns that have temporary immunity to the disease due to antibodies passed to them from their mother. Over time, this immunity will wane and they will move to the susceptible compartment [32, 51]. Once an individual has left the maternal immunity compartment, they cannot return to this compartment. Childhood diseases such as measles and poliomyelitis [52, 53] can be modelled using a compartment for maternal immunity.

- Treatment: A new compartment can be introduced for infected individuals who are undergoing treatment. This may reduce the duration that they are infected, reduce their infectiousness or both. Tuberculosis is an example of a disease that can be modelled with a treatment compartment [47].

- Vector compartments: Diseases such as malaria [54] and dengue fever [55] are transmitted by vectors (mosquitoes) rather than from person to person. To be able to model vector based diseases, additional compartments need to be added to model the disease within the vector population. A similar process can be used to model rabies [56] which is spread by dogs.

- Multiple infected compartments: For diseases where individuals remain infected for long periods of time or go through different stages of infection such as HIV and AIDS, multiple infected compartments can be added to the model [57]. Newly infected individuals start in the first compartment whereas those who have been infected longer will be in later infected compartments. This allows for each stage of infection to have different infectiousness and recovery times.

- Age based compartments: In the models that have been investigated in this chapter, it was assumed that the population was homogeneous. For example, each person within the population had the same chance of being infected and approximately the same recovery time. However, this is not always the case. Newborns and the elderly may be much more likely to contract the disease. They may also have longer recovery times. By splitting the population into different age groups, each age group can have their own values for the different parameters [58, 41]. For example, different contact rates (different $\beta$ values) can be used for different age compartments as well as different recovery rates.

- Spatial compartments: One of the assumptions made at the start of this chapter was that the entire population is well mixed. When modeling a disease over a very large area, this may not be the case. Therefore, it may be advantageous to split the population into different spatial compartments [59]. Only individuals within the same spatial compartment are now considered to be well mixed.
Transmission rates between different spatial compartments will generally be much lower than transmission rates within the same spatial compartment.

As more compartments are added to the model, the number of parameters will also rise. Unfortunately, many of the parameters for epidemic models are not known with certainty. This even applies to very important parameters such as the transmission rate. As the parameters are not known with certainty, adding extra compartments may not always lead to more accurate predictions.

## 2.2 Background on stochastic Galerkin method

In this section, the background to the Weiner-Hermite expansion is presented as well as information about generalised polynomial chaos. Both of these expansions form the foundation of the stochastic Galerkin method. The stochastic Galerkin method is also introduced and applied to a differential equation where one of the terms depends on a random variable.

### 2.2.1 Wiener-Hermite expansion

The Wiener-Hermite expansion is very important and forms the basis for generalised polynomial chaos. Introduced by Wiener in 1938 [60], it is also known as homogeneous chaos [61]. The Wiener-Hermite expansion is used to represent a second order random process (a process with finite variance) using Hermite polynomials and normally distributed (Gaussian) random variables.

The second order random process, $X(\omega)$, that depends on the outcome of a random event, $\omega$, can be written in the form [61]

$$X(\omega) = a_0 \Gamma_0 + \sum_{i_1=1}^{\infty} a_{i_1} \Gamma_1(\xi_{i_1}(\omega)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1,i_2} \Gamma_2(\xi_{i_1}(\omega), \xi_{i_2}(\omega)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \sum_{i_3=1}^{\infty} a_{i_1,i_2,i_3} \Gamma_3(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \xi_{i_3}(\omega)) + \ldots$$

where $\Gamma_n$ are the $n$th order, multi-dimensional, Hermite polynomials and $\xi_{i_j}$ ($j \in \{1, 2, 3, \ldots\}$) are independent, normally distributed, random variables with zero mean and unit variance.

The first few one-dimensional Hermite polynomials, $He_i(\xi_1)$, are given by [62]

$$He_0(\xi_1) = 1$$
$$He_1(\xi_1) = \xi_1$$
$$He_2(\xi_1) = \xi_1^2 - 1$$
$$He_3(\xi_1) = \xi_1^3 - 3\xi_1$$
and follow the orthogonality relationship [62]

\[ \int_{-\infty}^{\infty} H_i(\xi_1)H_j(\xi_1)w(\xi_1)d\xi_1 = n!\sqrt{2\pi}\delta_{ij} \]

where \( \delta_{ij} \) is the Kronecker delta and is defined as

\[ \delta_{ij} = \begin{cases} 
1 & \text{if } i = j \\
0 & \text{otherwise}.
\end{cases} \]

The weight function, \( w(\xi_1) \), of the Hermite polynomials is given by

\[ w(\xi_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi_1^2}{2}}. \]

It is important to note that the probability density function, \( p(\xi_1) \), of a one-dimensional, normally distributed random variable, \( \xi_1 \), is given by

\[ p(\xi_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\xi_1-\mu)^2}{2\sigma^2}}, \]

where \( \mu \) and \( \sigma \) are the mean and variance of \( \xi_1 \), respectively [62]. However, if the random variable \( \xi_1 \) has zero mean and unit variance, as with the normally distributed random variables associated with the Weiner-Hermite expansion, then its probability density function simplifies to

\[ p(\xi_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi_1^2}{2}}, \]

which matches the weight function of the Hermite polynomials.

As an example, consider \( X(\omega) \) to be a two-dimensional random process that relies on two independent random variables, \( \xi_1 \) and \( \xi_2 \), which are normally distributed with zero mean and unit variance. The Weiner-Hermite expansion gives

\[ X(\omega) = a_0\Gamma_0 + a_1\Gamma_1(\xi_1(\omega)) + a_2\Gamma_1(\xi_2(\omega)) + a_{1,1}\Gamma_2(\xi_1(\omega),\xi_1(\omega)) + a_{2,1}\Gamma_2(\xi_2(\omega),\xi_1(\omega)) + a_{2,2}\Gamma_2(\xi_2(\omega),\xi_2(\omega)) + \ldots \]

where \( \Gamma_n \) are the two-dimensional Hermite polynomials. However, for convenience and easier reading, a notation change can be introduced so that the Weiner-Hermite expansion can be rewritten as [61]

\[ X(\omega) = \hat{a}_0\Psi_0(\xi) + \hat{a}_1\Psi_1(\xi) + \hat{a}_2\Psi_2(\xi) + \hat{a}_3\Psi_3(\xi) + \hat{a}_4\Psi_4(\xi) + \hat{a}_5\Psi_5(\xi) + \ldots \]
where, for example,
\[
\hat{a}_0 = a_0
\]
\[
\hat{a}_1 = a_1
\]
\[
\hat{a}_2 = a_2
\]
\[
\hat{a}_3 = a_{1,1}
\]
\[
\hat{a}_4 = a_{2,1}
\]
\[
\hat{a}_5 = a_{2,2}
\]

and, similarly,
\[
\Psi_0(\xi) = \Gamma_0
\]
\[
\Psi_1(\xi) = \Gamma_1(\xi_1(\omega))
\]
\[
\Psi_2(\xi) = \Gamma_1(\xi_2(\omega))
\]
\[
\Psi_3(\xi) = \Gamma_2(\xi_1(\omega),\xi_1(\omega))
\]
\[
\Psi_4(\xi) = \Gamma_2(\xi_2(\omega),\xi_1(\omega))
\]
\[
\Psi_5(\xi) = \Gamma_2(\xi_2(\omega),\xi_2(\omega))
\]

where $\xi = \{\xi_1(\omega),\xi_2(\omega)\}$. Therefore the Weiner-Hermite expansion can be written using a single summation [25] and is given by
\[
X(\omega) = \sum_{i=0}^{\infty} \hat{a}_i \Psi_i(\xi).
\] (2.26)

While this makes the Weiner-Hermite expansion much easier to read and write, it is important to note that when written in this form, the order of Hermite polynomial, $\Psi_i(\xi)$, is not necessarily given by its subscript, $i$. For example, $\Psi_5(\xi)$ is actually a second order Hermite polynomial, as can be seen above.

### 2.2.2 Generalised polynomial chaos

Generalised polynomial chaos (gPC) was designed to be an extension to the Weiner-Hermite expansion (Equation 2.26) [63]. While the Weiner-Hermite expansion is useful when the inputs are normally distributed, its convergence is slow when using inputs that are not normally distributed [64, 21]. Rather than always using Hermite polynomials, generalised polynomial chaos (or Weiner-Askey polynomial chaos) uses other orthogonal polynomials from the Askey scheme [65, 66]. The type of orthogonal polynomial is carefully chosen such that the probability density function of the random variables matches the weight function of the orthogonal polynomials. Sim-
ilar to the Weiner-Hermite expansion, the generalised polynomial chaos expansion is given by

\[ X(\omega) = \sum_{i=0}^{\infty} \hat{a}_i \Phi_i(\xi). \]  

(2.27)

where \( \Phi_i(\xi) \) are appropriately chosen orthogonal polynomials from the Askey scheme.

Table 2.1 shows the orthogonal polynomials associated with the most common probability distributions [21]. The probability density function (PDF) as well as the support of the random variables is also given in the table. Note that the support of the random variable is also the domain of the orthogonal polynomials. As was discussed with the Weiner-Hermite expansion, when the random variable is normally distributed with zero mean and unit variance, the Hermite polynomials should be used as basis functions. If the random variable is normally distributed but does not have zero mean or unit variance, its probability density function will not match the weight function of the Hermite polynomials.

When the random variable is uniformly distributed on \([-1, 1]\), the associated orthogonal polynomials are the Legendre polynomials. While it is possible to have a uniform distribution on any compact domain, in order to use the Legendre polynomials, which are defined on \([-1, 1]\), the uniform distribution must also be on \([-1, 1]\). Uniform distributions on other domains will be discussed in Chapter 3. It does not matter that the probability density function of the uniformly distributed random variable and the weight function of the Legendre polynomials differ by a constant factor.

If the random variable is beta distributed, the associated orthogonal polynomials are the Jacobi polynomials. Note that this is not the ‘standard’ beta distribution, which is most commonly defined on \([0,1]\), but instead the generalised beta distribution on \([-1,1]\) [29]. The generalised beta distribution is used so that its probability density function matches the weight function of the Jacobi polynomials. Note that if \(\alpha = 0\) and \(\beta = 0\) are used in the generalised beta distribution, it simplifies to a uniform distribution on \([-1,1]\). Similarly, the weight function of the Jacobi polynomials simplifies to 1, which is the weight function of the Legendre polynomials.

When the random variable is gamma distributed with unit shape and scale parameter, the associated orthogonal polynomials are the Laguerre polynomials. It is important for the the gamma distributed random variable to have unit shape and scale parameters so that the probability density function of the random variable matches the weight function of the Laguerre polynomials.

Using generalised polynomial chaos, rather than the Weiner-Hermite expansion, allows for the input random variables to have a wider range of probability distributions.
Table 2.1: A few common distributions and their associated orthogonal polynomials from the Askey scheme. Note that $B$ is the beta function where $B(\alpha + 1, \beta + 1) = \frac{\alpha!\beta!}{(\alpha + \beta + 1)!}$.

### 2.2.3 Stochastic Galerkin method

The stochastic Galerkin method is an extension to the commonly used Galerkin method which is used to solve deterministic differential equations. The stochastic Galerkin method uses the generalised polynomial expansion (Equation 2.27) to solve differential equations that rely on random inputs [64, 21]. A short explanation of the stochastic Galerkin method in one dimension is given below, but a much more detailed analysis of the stochastic Galerkin method is conducted in Chapter 3 when it is applied to epidemic models.

Consider a linear differential operator, $L$, and the deterministic differential equation given by

$$L[u(x)] = f(x),$$

where $f(x)$ is a known function and $u(x)$ is the solution to the differential equation. Depending on the form of the linear differential operator, $L$, that is used in this differential equation, there are numerical methods, and possibly even analytical methods, that can be used to solve for $u(x)$. However, rather than $f(x)$ being a deterministic function, instead consider it to rely on the outcome of a random variable, $\xi$, which has the probability density function $w(\xi)$. As $f$ is a function of $x$ as well as the random variable, it can now be written as $f(x, \xi)$. Also, as $f$ depends upon a random variable, the solution $u$ will also rely on the random variable and is now written as $u(x, \xi)$. Therefore the differential equation is now

$$L[u(x, \xi)] = f(x, \xi).$$

Using the generalised polynomial expansion, $u(x, \xi)$ and $f(x, \xi)$ can be expanded in the form

$$u(x, \xi) = \sum_{i=0}^{\infty} u_i(x) \Phi_i(\xi),$$

$$f(x, \xi) = \sum_{i=0}^{\infty} f_i(x) \Phi_i(\xi),$$

where $\Phi_i(\xi)$ are orthogonal polynomials.
where \( u_i(x) \) and \( f_i(x) \) are deterministic functions and \( \Phi_i(\xi) \) are appropriately chosen orthogonal polynomials. The orthogonal polynomials are chosen such that the weight function of the polynomials matches the probability density function of the random variable (see Table 2.1).

It is important to note that the generalised polynomial chaos expansions are comprised of constants, \( \hat{a}_i \), and orthogonal polynomials, \( \Phi_i \) (see Equation (2.27)). However, the stochastic Galerkin expansions are comprised of the functions \( u_i(x) \) and \( f_i(x) \), as well as the orthogonal polynomials \( \Phi_i(\xi) \). Therefore, to use the polynomial chaos expansions, only the value of the constants must be found. However, when using the stochastic Galerkin method, deterministic functions need to be calculated.

As this is a numerical solution, the expansions will need to be truncated after an appropriate number of terms and are now given by

\[
    u(x, \xi) \approx \sum_{i=0}^{P} u_i(x) \Phi_i(\xi) \quad (2.30)
\]

\[
    f(x, \xi) \approx \sum_{i=0}^{P} f_i(x) \Phi_i(\xi), \quad (2.31)
\]

where the expansions have been truncated at \( P + 1 \) terms. Substituting Equations (2.30) and (2.31) into Equation (2.28) gives

\[
    L \left[ \sum_{i=0}^{P} u_i(x) \Phi_i(\xi) \right] = \sum_{i=0}^{P} f_i(x) \Phi_i(\xi),
\]

which can be simplified to

\[
    \sum_{i=0}^{P} L [u_i(x)] \Phi_i(\xi) = \sum_{i=0}^{P} f_i(x) \Phi_i(\xi), \quad (2.32)
\]

as \( \Phi_i(\xi) \) are only functions of the random variable \( \xi \).

To calculate the deterministic functions, \( f_i(x) \), a Galerkin projection is applied to Equation (2.31) by first multiplying by \( \Phi_j(\xi) \), \( (j = 0, 1, \ldots, P) \), and integrating over the probability space of the random variable \( \xi \). This gives

\[
    \langle f(x, \xi), \Phi_j(\xi) \rangle = \sum_{i=0}^{P} f_i(x) \langle \Phi_i(\xi), \Phi_j(\xi) \rangle,
\]

where the inner product, \( \langle \Phi_i(\xi), \Phi_j(\xi) \rangle \), is defined as

\[
    \langle \Phi_i(\xi), \Phi_j(\xi) \rangle = \int_{\Omega} \Phi_i(\xi) \Phi_j(\xi) w(\xi) d\xi,
\]

and \( \Omega \) is the sample space of \( \xi \) (the set of all possible outcomes). Also, recall that \( w(\xi) \) is the probability density function of \( \xi \).
As the orthogonal polynomials were carefully chosen such that their weight function matches the probability density function of the random variable, many of the inner products trivially evaluate to zero as
\[
\langle \Phi_i(\xi), \Phi_j(\xi) \rangle = \begin{cases} 
0 & \text{if } i \neq j \\
\langle \Phi_i(\xi), \Phi_i(\xi) \rangle & \text{otherwise.}
\end{cases}
\]
Therefore the deterministic functions \( f_i(x) \) are given by
\[
f_i(x) = \frac{\langle f(x, \xi), \Phi_i(\xi) \rangle}{\langle \Phi_i(\xi), \Phi_i(\xi) \rangle},
\]
for \( i = 0, 1, \ldots, P \). Once the functions \( f_i(x) \) have been calculated, the functions \( u_i(x) \) can be calculated. In a similar fashion to previously, this is achieved by applying a Galerkin projection to Equation (2.32) which gives
\[
\sum_{i=0}^{P} L \left[u_i(x)\right] \langle \Phi_i(\xi), \Phi_j(\xi) \rangle = \sum_{i=0}^{P} f_i(x) \langle \Phi_i(\xi), \Phi_j(\xi) \rangle.
\]
(2.33)
Once again, as \( \Phi_i(\xi) \) are orthogonal polynomials, many of the inner products trivially evaluate to zero. This leaves a system of \( P + 1 \) deterministic differential equations of the form
\[
L[u_i(x)] = f_i(x) = \frac{\langle f(x, \xi), \Phi_i(\xi) \rangle}{\langle \Phi_i(\xi), \Phi_i(\xi) \rangle},
\]
for \( i = 0, 1, \ldots, P \). It is important to note that while this example started with a single differential equation with a random input, and ended as a system of \( P + 1 \) differential equations, the final system of equations in the stochastic Galerkin method is deterministic and only needs to be solved once. An analysis of computational cost of the stochastic Galerkin method and its accuracy is presented in Chapter 3, as well as the situation where there are multiple random inputs.
Chapter 3

Applying the stochastic Galerkin method to the $SIR$ model

Mathematically, it would seem as though epidemic compartment models, as derived in the previous chapter, would work quite well in predicting the likely course of an epidemic. However, many of the parameters used within these compartment models are not known with certainty and have to be estimated from available data. As the data available is not always complete and accurate, this can lead to very different parameter estimates [67].

For example, the number of people an individual comes into contact with during their day is going to differ greatly from person to person which will affect estimates for the transmission coefficient, $\beta$. Some possible explanations for an individual’s differing number of contacts within the population include:

- where they live within the population (city versus suburbs)
- where they work (a small office or business with only a few workers and customers versus a large corporate office building, large factory with many workers or in a shopping complex with many customers daily)
- mode of transportation (car versus public transport)
- social habits (socialise with small groups of friends versus large parties).

Even for a single person within the population, the number of contacts they make is going to differ from day to day. This makes estimates for the transmission coefficient, $\beta$, quite difficult.

The recovery rate, $1/\gamma$, can also be difficult to measure and can often lead to different parameter estimates as individuals infected with the same disease can recover at different rates [68, 69].

Initial conditions, such as the initial number of infected individuals, can also be difficult to determine. While it is often assumed that there are no recovered individuals, $R$, at the start of the epidemic, this may not be the case. The number
Chapter 3. Applying the stochastic Galerkin method to the SIR model

of individuals that are immune to the disease, through vaccination [9] or having the
disease previously, can be extremely difficult to determine.

As the parameter values cannot be determined exactly, their estimates contain
some degree of uncertainty. It is important for this uncertainty to be included into
the epidemic model, or the model may give misleading predictions.

In this chapter, SIR models with different types of uncertainty will be inves-
tigated. The mean predictions will then be calculated using both Monte Carlo
sampling as well as the stochastic Galerkin method. The results from the two differ-
ent methods will then be compared. Random variables with uniform, gamma and
normal distributions will be considered.

3.1 Representing uncertainty in parameters and
initial conditions

In order to represent the uncertainty in the parameters of an epidemic model, rather
than assuming the parameters are constants, they are instead considered functions of
random variables [12]. For example, the most common parameters with uncertainty,
$\beta$ and $\gamma$, can be written in the form

$$
\beta = \beta_0 + a \xi_1 \\
\gamma = \gamma_0 + b \xi_2,
$$

where $\beta_0$ and $\gamma_0$ are positive constants, $\xi_1$ and $\xi_2$ are continuous random variables
with known probability distributions, and $a$ and $b$ are positive scaling constants.
It is important to note that $\xi_1$ and $\xi_2$ may not have zero mean. Whereas in the
previous chapter $\beta$ and $\gamma$ were constants, they are now probability distributions.
Writing uncertain parameters in this form is not just restricted to $\beta$ and $\gamma$. Any
parameter that includes uncertainty can be expressed this way.

For new diseases where little is known about the parameters, wide probability
distributions can be used to represent the limited knowledge of the parameter values.
For common and seasonal diseases where the values of the parameters have been
widely studied and the values of the parameters are known with some certainty,
narrow probability distributions can be used to represent the relative certainty in
the values of the parameters.

Similarly, random variables can also be used to represent uncertainty in the
initial conditions. For example, the initial fraction of infected individuals, $I_{\text{initial}}$,
can be rewritten as

$$
I_{\text{initial}} = C + g \xi_3,
$$

where $C$ and $g$ are positive constants and $\xi_3$ is a continuous random variable with
a known probability distribution.
Chapter 3. Applying the stochastic Galerkin method to the SIR model

3.2 SIR model with uncertainty in the transmission rate (\(\beta\))

Now that the uncertainty in the parameters and initial conditions of an epidemic model can be introduced into the model by using random variables, the effect that this has on an SIR model can be investigated. For simplicity, in this first example, it is assumed that there is only uncertainty in the transmission rate, \(\beta\), and that the recovery rate and initial conditions are known with certainty. The equations for an SIR model (see Section 2.1.3 and Equations (2.12)-(2.14)) with uncertainty in the transmission rate are given by

\[
\begin{align*}
\frac{dS}{dt} &= -(\beta_0 + a\xi)SI \\
\frac{dI}{dt} &= (\beta_0 + a\xi)SI - \gamma I \\
\frac{dR}{dt} &= \gamma I,
\end{align*}
\]

where \(\beta_0\) and \(a\) are positive constants and \(\xi\) is a random variable with a known probability distribution. Now that there is uncertainty in the transmission rate, \(\beta\) in the deterministic SIR model has been replaced with \(\beta_0 + a\xi\). Note that the differential equation for \(R\) is not needed as \(R\) is simply given by \(R = 1 - S - I\). This example is fundamentally the same as the model investigated by Roberts [12]. A discussion of Roberts’ paper is presented in Section 3.4.

As there is no known analytical solution for the SIR model, even the deterministic SIR model needs to be solved numerically. While this is usually a relatively simple process, this SIR model contains a random variable. Therefore, the model can no longer be solved using a single function call to a numerical ODE solver.

By generating a value for \(\xi\) based upon its probability distribution, a single realisation of the model can be obtained using a numerical ODE solver. Additional values for \(\xi\) can then be generated to determine other possible realisations of the model. Depending on the degree of uncertainty in \(\beta\) (and therefore the probability distribution of \(\xi\)), this could lead to very different predictions.

By generating a large number of realisations, the mean, variance and other statistical moments of the predictions from the SIR model can be determined. This process is known as Monte Carlo sampling [70]. Assuming that a large number of realisations is used, Monte Carlo sampling provides accurate results. However, it can be very computationally expensive due to the large number of times the model must be solved for the Monte Carlo solution to converge.

Figure 3.1 shows an example prediction for an SIR model with uncertainty in the transmission rate. The figure shows the mean prediction as well as its variance when \(\beta_0 = 4\), \(\gamma = 1\), \(a = 1\) and \(\xi\) has a uniform distribution on \([-1, 1]\). The
Figure 3.1: Solutions for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. Solid lines calculated using Monte Carlo sampling ($a = 1$) while dashed lines use average value of $\beta$ ($a = 0$). $\beta_0 = 4$, $\gamma = 1$, $\xi$ is a continuous random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. 
probability distribution for $\beta$ can be seen in Figure 3.2(a). It is important to note that despite the uncertainty in $\beta$, $R_0 > 1$ for all possible values of $\xi$. The results in Figure 3.1 were obtained using Monte Carlo sampling with $10^5$ realisations. The mean and variance were also calculated using $2 \times 10^5$ realisations and compared to those obtained using $10^5$ realisations to ensure that $10^5$ realisations was sufficient to accurately calculate the mean and variance. The maximum absolute difference between the mean predictions for $S$ and $I$ was less than $1.5 \times 10^{-3}$ for all $t \in [0, 6]$. Also, the maximum absolute difference between the variance for $S$ and $I$ was less than $5 \times 10^{-4}$ for all $t \in [0, 6]$. Therefore, by using such a large number of Monte Carlo trials ($10^5$), it can be assumed that the predictions obtained are very close to the exact mean and variance.

Along with the predictions from Monte Carlo sampling, the prediction using only the ‘average’ value of $\beta$ (equivalent to $a = 0$) is also shown in Figure 3.1(a). From the graph it can be seen that the two predictions are different, with a slightly higher peak in the epidemic curve when only the average value of $\beta$ is used. From the graph of the variance (Figure 3.1(b)), it can be seen that at the start of the epidemic ($t = 0$), the variance starts at zero as there was no uncertainty in the initial conditions. The variance for both the susceptible and infected populations then increases to reach its maximum, which occurs at approximately the same time as the peak in the epidemic. The variance of both the susceptible and infected populations then decreases. The variance in the infected population reaches almost zero by approximately $t = 2.3$ before increasing again slightly, for a second smaller peak, before returning to almost zero. The variance in the susceptible population does not have this second peak. Instead, it simply decreases from its peak to almost zero. When using only the average value of $\beta$ ($a = 0$), the variance is zero as it becomes a deterministic model.

From Figure 3.1, it can be seen that it is important to incorporate uncertainty into the model, as not including the uncertainty leads to a different prediction. This
can be more clearly seen in Figure 3.3 where \( a \) has been increased from 1 to 2, which effectively doubles the uncertainty in \( \beta \). The new probability distribution for \( \beta \) can be seen in Figure 3.2(b). Despite the increased uncertainty in \( \beta \), \( R_0 > 1 \) for all possible values of \( \beta \). The predictions obtained from Monte Carlo sampling as well as the deterministic model \( (a = 0) \) are significantly different. Also, not surprisingly, now that the uncertainty in \( \beta \) has been increased, the magnitude of the variance has also significantly increased. The variance still starts at zero as there was no uncertainty in the initial conditions, again reaches its peak at approximately the same time as the peak in the epidemic, and then decreases to almost zero. The effect of uncertainty in the initial conditions, as well as the recovery rate, will be investigated in later sections of this chapter.

### 3.3 Using the stochastic Galerkin method

While accurate results for the mean and variance can be obtained using Monte Carlo sampling (assuming a large number of trials is used), it can be very computationally expensive. The computational cost of the Monte Carlo method, as well as its accuracy using different number of trials, is investigated later in this chapter. Also, as it is an entirely numerical approach, if any of the values of the parameters are changed, all of the Monte Carlo sampling must be redone using the new parameter values as previous calculations are unlikely to be helpful in determining the new mean and variance.

Rather than using a ‘brute force’ Monte Carlo sampling technique to obtain the mean solution and its variance, the stochastic Galerkin method can instead be used [64]. The stochastic Galerkin method assumes that the solution to the SIR model (Equations (3.1)) can be written in the form [61, 25]

\[
S(t, \xi) = \sum_{i=0}^{\infty} S_i(t) \Psi_i(\xi) \\
I(t, \xi) = \sum_{i=0}^{\infty} I_i(t) \Psi_i(\xi),
\]

(3.2)

where \( S_i(t) \) and \( I_i(t) \) are deterministic functions (that need to be determined) and \( \Psi_i(\xi) \) are appropriately chosen orthogonal polynomials. The orthogonal polynomials, \( \Psi_i(\xi) \), form a basis over which the solutions can be expanded (similar to the finite element method). It is important to note that, whereas before \( S \) and \( I \) were only functions of time (when uncertainty was not incorporated into the model), it is now assumed that they are functions not only of time, but functions of the random variable, \( \xi \), as well.

While, analytically, the solution is the sum of an infinite number of terms, in practice the series needs to be truncated at an appropriate point. It is expected
Figure 3.3: Solutions for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. Solid lines calculated using Monte Carlo sampling ($a = 2$) while dashed lines use average value of $\beta$ ($a = 0$). $\beta_0 = 4$, $\gamma = 1$, $\xi$ is a continuous random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. 

$\beta$
that each additional term is smaller in magnitude than the previous term so the series can be truncated while still giving an accurate solution [71]. The accuracy of different order expansions will be investigated later in this chapter. Truncating the series at $P + 1$ terms gives

$$S(t, \xi) = \sum_{i=0}^{\infty} S_i(t) \Psi_i(\xi) \approx \sum_{i=0}^{P} S_i(t) \Psi_i(\xi)$$

$$I(t, \xi) = \sum_{i=0}^{\infty} I_i(t) \Psi_i(\xi) \approx \sum_{i=0}^{P} I_i(t) \Psi_i(\xi).$$

Substituting the new forms of the solution (Equations (3.3)) into the $SIR$ model (Equations (3.1)) gives

$$\sum_{i=0}^{P} \frac{dS_i(t)}{dt} \Psi_i(\xi) = -(\beta_0 + a \xi) \sum_{i=0}^{P} S_i(t) \Psi_i(\xi) \sum_{j=0}^{P} I_j(t) \Psi_j(\xi)$$

$$\sum_{i=0}^{P} \frac{dI_i(t)}{dt} \Psi_i(\xi) = (\beta_0 + a \xi) \sum_{i=0}^{P} S_i(t) \Psi_i(\xi) \sum_{j=0}^{P} I_j(t) \Psi_j(\xi)$$

$$- \gamma \sum_{i=0}^{P} I_i(t) \Psi_i(\xi). \quad (3.4)$$

The next step is to perform a Galerkin projection. To perform a Galerkin projection, the equations are first multiplied through by $\Psi_k(\xi)$ ($k = 0, 1, \ldots, P$). The equations are then integrated over the probability space. This is done by first multiplying through by $w(\xi)$, where $w(\xi)$ is the probability density function for the random variable $\xi$. A definite integral is then performed on the equations, with respect to the random variable, with the limits of integration being the minimum and maximum values of the random variable $\xi$. For example, if $\xi$ is a random variable with a uniform distribution on $[-1, 1]$, then $w(\xi) = 1/2$ and the limits of integration would be $-1$ and $1$. Therefore, after performing a Galerkin projection, Equations (3.4) become

$$\sum_{i=0}^{P} \frac{dS_i(t)}{dt} \langle \Psi_i(\xi), \Psi_k(\xi) \rangle = -\beta_0 \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t) I_j(t) \langle \Psi_i(\xi) \Psi_j(\xi), \Psi_k(\xi) \rangle$$

$$- a \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t) I_j(t) \langle \xi \Psi_i(\xi) \Psi_j(\xi), \Psi_k(\xi) \rangle \quad (3.5)$$
Chapter 3. Applying the stochastic Galerkin method to the SIR model

\[
\sum_{i=0}^{P} \frac{dI_i(t)}{dt} \langle \Psi_i(\xi), \Psi_k(\xi) \rangle = \beta_0 \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t) I_j(t) \langle \Psi_i(\xi) \Psi_j(\xi), \Psi_k(\xi) \rangle \\
+ a \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t) I_j(t) \langle \xi \Psi_i(\xi) \Psi_j(\xi), \Psi_k(\xi) \rangle \\
- \gamma \sum_{i=0}^{P} I_i(t) \langle \Psi_i(\xi), \Psi_k(\xi) \rangle
\]

where

\[
\langle \Psi_i(\xi), \Psi_k(\xi) \rangle = \int_{\Omega} \Psi_i(\xi) \Psi_k(\xi) w(\xi) d\xi
\]

and \( \Omega \) is the sample space of \( \xi \) (the set of all all possible outcomes).

From Equations (3.5) and (3.6) it is easy to see why orthogonal polynomials are chosen as basis functions for the stochastic Galerkin method. If \( \Psi_i(\xi) \) are orthogonal polynomials with weight function \( w(\xi) \) (ie the weight function of the orthogonal polynomials matches the probability density function of the random variable), many of the inner products trivially evaluate to zero.

For example, if \( \xi \) is a uniform random variable on \([-1,1]\], the Legendre polynomials would be chosen as basis functions for the stochastic Galerkin expansion [64]. The first few Legendre polynomials, \( P_i(\xi) \), are given by [62]

\[
P_0(\xi) = 1 \\
P_1(\xi) = \xi \\
P_2(\xi) = (3(\xi)^2 - 1)/2 \\
P_3(\xi) = (5(\xi)^3 - 3\xi)/2.
\]

Higher order Legendre polynomials can quickly be calculated using the recurrence relation [62]

\[
(i + 1) P_{i+1}(\xi) = (2i + 1) \xi P_i(\xi) - iP_{i-1}(\xi).
\]

By evaluating the inner products in Equations (3.5) and (3.6), a deterministic system of \( 2(P + 1) \) simultaneous ordinary differential equations can be derived. For example, the final deterministic system of ODEs obtained using a second order stochastic Galerkin expansion \( (P = 2) \) when \( \xi \) is a uniformly distributed random variable on \([-1,1]\) is given by

\[
\frac{dS_0}{dt} = -\beta_0(S_0 I_0 + \frac{S_1 I_1}{3} + \frac{S_2 I_2}{5}) \\
- a(\frac{S_0 I_1}{3} + \frac{S_1 I_0}{3} + \frac{2S_1 I_2}{15} + \frac{2S_2 I_1}{15})
\]

\[
\frac{dS_1}{dt} = -\beta_0(S_0 I_1 + S_1 I_0 + \frac{2S_1 I_2}{5} + \frac{2S_2 I_1}{5}) \\
- a(S_0 I_0 + \frac{2S_0 I_2}{5} + \frac{3S_1 I_1}{5} + \frac{2S_2 I_0}{5} + \frac{11S_2 I_2}{35})
\]

(3.7)
\[
\frac{dS_2}{dt} = -\beta_0 \left( S_0 I_2 + \frac{2S_1 I_1}{3} + S_2 I_0 + \frac{2S_2 I_2}{7} \right) - a \left( \frac{2S_0 I_1}{3} + \frac{2S_1 I_0}{3} + \frac{11S_1 I_2}{21} + \frac{11S_2 I_1}{21} \right) \tag{3.9}
\]
\[
\frac{dI_0}{dt} = \beta_0 \left( S_0 I_0 + \frac{S_1 I_1}{3} + \frac{S_2 I_2}{5} \right) + a \left( \frac{S_0 I_1}{3} + \frac{S_1 I_0}{3} + \frac{2S_1 I_2}{15} + \frac{2S_2 I_1}{15} \right) - \gamma I_0 \tag{3.10}
\]
\[
\frac{dI_1}{dt} = \beta_0 \left( S_0 I_1 + S_1 I_0 + \frac{2S_1 I_2}{5} + \frac{2S_2 I_1}{5} \right) + a \left( S_0 I_0 + \frac{2S_0 I_2}{5} + \frac{3S_1 I_1}{5} + \frac{2S_2 I_0}{5} + \frac{11S_2 I_2}{35} \right) - \gamma I_1 \tag{3.11}
\]
\[
\frac{dI_2}{dt} = \beta_0 \left( S_0 I_2 + \frac{2S_1 I_1}{3} + S_2 I_0 + \frac{2S_2 I_2}{7} \right) + a \left( \frac{2S_0 I_1}{3} + \frac{2S_1 I_0}{3} + \frac{11S_1 I_2}{21} + \frac{11S_2 I_1}{21} \right) - \gamma I_2. \tag{3.12}
\]

After performing the Galerkin projection, the random variable, \( \xi \), no longer appears in the final system of equations. This deterministic system of differential equations can easily be solved using a numerical ODE solver such as the MATLAB function \texttt{ode45}. While uncertainty was explicitly included into the epidemic model (by making \( \beta \) a function of a random variable), the random variable itself no longer appears in the final system of equations.

In order to solve Equations (3.7)-(3.12) numerically, initial conditions for \( S_i(t) \) and \( I_i(t) \) first need to be determined. In this example, it was assumed that there was no uncertainty in the initial conditions and that there were initially \( I_{\text{initial}} \) infected, where \( 0 < I_{\text{initial}} \leq 1 \). The remaining population was susceptible to the disease. Therefore, \( S(0, \xi) = 1 - I_{\text{initial}} \) and \( I(0, \xi) = I_{\text{initial}} \).

In terms of the stochastic Galerkin expansion, the initial condition for the fraction infected is given by

\[
I(0, \xi) = I_{\text{initial}} = \sum_{i=0}^{P} I_i(0) \Psi_i(\xi). 
\]

Performing a Galerkin projection on the initial condition by multiplying by \( \Psi_k(\xi) \) and integrating over the probability space gives

\[
I_{\text{initial}} \int_{\Omega} \Psi_k(\xi) w(\xi) d\xi = \sum_{i=0}^{P} I_i(0) \int_{\Omega} \Psi_i(\xi) \Psi_k(\xi) w(\xi) d\xi. 
\]

However, as

\[
\int_{\Omega} \Psi_k(\xi) w(\xi) d\xi = \begin{cases} 
1 & \text{if } k = 0 \\
0 & \text{if } k \geq 1
\end{cases} 
\]

because of the orthogonality of the Legendre polynomials, the initial conditions for
$I_i(t)$ are simply given by

$$I_i(0) = \begin{cases} I_{\text{initial}} & \text{if } i = 0 \\ 0 & \text{if } i > 0. \end{cases}$$

Similarly, the initial conditions for $S_i(t)$ are given by

$$S_i(0) = \begin{cases} 1 - I_{\text{initial}} & \text{if } i = 0 \\ 0 & \text{if } i > 0. \end{cases}$$

Figure 3.4 shows the solutions for $S_i(t)$ and $I_i(t)$ obtained from an ODE solver for $P = 5$ (which is a slightly larger system of equations than Equations (3.7)-(3.12). It is interesting to note that while $S$ and $I$ cannot be negative, some of the solutions for $S_i(t)$ and $I_i(t)$ are negative. Also, as $i$ increases, the solutions for $S_i(t)$ and $I_i(t)$ decrease in magnitude. For example, $S_0(t)$ goes up to almost 1 and $S_1(t)$ goes down to approximately -0.5, however $S_5(t)$ stays much closer to zero. As the magnitude of the higher order terms is much smaller than the lower order terms, it makes sense that the expansions can be truncated at $P + 1$ terms while still producing accurate results.

### 3.3.1 Determining mean and variance from the stochastic Galerkin solution

Now that the stochastic Galerkin solution has been determined, the next step is to determine the mean and variance from the stochastic Galerkin solution. The stochastic Galerkin solution for the fraction of infected individuals is given by

$$I(t, \xi) = \sum_{i=0}^{P} I_i(t) \Psi_i(\xi),$$

where $\xi$ is a continuous random variable with probability density function $w(\xi)$ and sample space $\Omega$.

The mean fraction of infected individuals $E[I(t, \xi)]$ is therefore given by [25]

$$E[I(t, \xi)] = \int_{\Omega} \sum_{i=0}^{P} I_i(t) \Psi_i(\xi) w(\xi) d\xi = \sum_{i=0}^{P} I_i(t) \int_{\Omega} \Psi_i(\xi) w(\xi) d\xi = I_0(t).$$

Therefore the mean solution is simply the zero order term of the stochastic Galerkin expansion. It is important to note that while the mean is simply given by the zero order term, it is not sufficient to simply derive a zero order expansion ($P = 0$) to obtain the mean solution. From Equations (3.7)-(3.12), it is easy to see that the deterministic ordinary differential equations are coupled and not deriving higher...
Figure 3.4: Solutions of $S_i(t)$ and $I_i(t)$ for an SIR model with uncertainty in $\beta$. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 2$, $\xi$ is a random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. 
order terms affects the lower order terms. However, as the order of the stochastic Galerkin expansion increases, the higher order terms have less of an effect on the lower order solutions.

The variance in the number of infected, \( \text{Var}[I(t, \xi)] \), is given by [25]

\[
\text{Var}[I(t, \xi)] = E[(I(t, \xi))^2] - (E[I(t, \xi)])^2
\]

\[
= \int_\Omega \left( \sum_{i=0}^{P} I_i(t)\Psi_i(\xi) \sum_{j=0}^{P} I_j(t)\Psi_j(\xi) \right) w(\xi) d\xi - (I_0(t))^2
\]

\[
= \sum_{i=0}^{P} \sum_{j=0}^{P} I_i(t) I_j(t) \int_\Omega \Psi_i(\xi)\Psi_j(\xi) w(\xi) d\xi - (I_0(t))^2.
\]

However, as \( \Psi_i(\xi) \) are orthogonal polynomials, the integral will evaluate to zero when \( i \neq j \) giving

\[
\text{Var}[I(t, \xi)] = \sum_{i=0}^{P} (I_i(t))^2 \int_\Omega \Psi_i(\xi)\Psi_i(\xi) w(\xi) d\xi - (I_0(t))^2
\]

\[
= \sum_{i=1}^{P} (I_i(t))^2 \langle \Psi_i(\xi), \Psi_i(\xi) \rangle. \quad (3.13)
\]

Further, if the orthogonal polynomials are normalised such that

\[
\int_\Omega \Psi_i(\xi)\Psi_i(\xi) w(\xi) d\xi = 1,
\]

the expression for the variance simplifies to

\[
\text{Var}[I(t, \xi)] = \sum_{i=1}^{P} (I_i(t))^2.
\]

Therefore, the variance is simply the sum of the square of the non-zero order terms of the stochastic Galerkin solution. By determining the stochastic Galerkin solution, both the mean and its variance can be quickly calculated without needing random sampling techniques such as Monte Carlo sampling.

However, it is important to note that not all orthogonal polynomials are normalised. For example, the Legendre polynomials, \( P_i(\xi) \), which are used if the random variable has a uniform distribution on \([-1,1]\), are not normalised and instead have the orthogonality relationship

\[
\int_{-1}^{1} P_i(\xi)P_j(\xi) \frac{1}{2} d\xi = \begin{cases} 
0 & \text{if } i \neq j \\
\frac{1}{2i+1} & \text{if } i = j.
\end{cases}
\]

Therefore, if the orthogonal polynomials are not normalised, Equation (3.13) must be used to calculate the variance.
Figure 3.5: Comparison of stochastic Galerkin solution with Monte Carlo sampling for SIR model with uncertainty in $\beta$. Blue is susceptible and red is infected. Solid lines are Monte Carlo solutions and dashed lines are stochastic Galerkin solutions. $\beta_0 = 4$, $\gamma = 1$, $a = 2$, $\xi$ is a random variable with a uniform distribution on $[-1, 1]$, $I_{\text{initial}} = 0.01$. 
Figure 3.5 shows a comparison of the solutions obtained from Monte Carlo sampling (with $10^5$ trials) and the stochastic Galerkin method, with the stochastic Galerkin solution being truncated at the first, third and fifth order expansions ($P = 1$, $P = 3$ and $P = 5$, respectively). Due to the large number of Monte Carlo trials, the Monte Carlo solution will be very close the exact solution and can be used to evaluate the accuracy of the stochastic Galerkin solution.

Even using a very low order expansion, $P = 1$, the mean solution obtained from the stochastic Galerkin method is very close to the ‘exact’ solution obtained from Monte Carlo sampling. However, the variance obtained from the stochastic Galerkin solution is significantly different to the variance of the ‘exact’ solution obtained from Monte Carlo sampling. As the order of the stochastic Galerkin expansion is increased, it converges to the solution obtained through Monte Carlo sampling. The fifth order stochastic Galerkin expansion is almost identical to that obtained through Monte Carlo sampling. This shows that when using the stochastic Galerkin method, truncated expansions still produce accurate results and it is not necessary to use high order expansions, for example $P > 10$, to obtain accurate results. This is reinforced in later sections of this chapter as low order expansions still provide accurate results when the uncertain parameters no longer have uniform distributions.

3.4 Discussion of Roberts’ paper

In his paper, *Epidemic models with uncertainty in the reproduction number* [12], Roberts investigates an SIR model with uncertainty in the reproduction number. This section discusses the work already done by Roberts in applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters, as well as the limitations of his approach.

To begin, the equations used in Roberts’ paper will be derived. (Note that some of the notation has been changed to remain consistent with this thesis). Starting with a deterministic SIR model given by

\[
\frac{dS}{dt} = -\beta SI \tag{3.14}
\]
\[
\frac{dI}{dt} = \beta SI - \gamma I, \tag{3.15}
\]

Roberts then normalised time with respect to the recovery rate ($\gamma$) such that $\tau = \gamma t$ giving

\[
\frac{dS}{d\tau} = -R_0 SI \tag{3.16}
\]
\[
\frac{dI}{d\tau} = R_0 SI - I, \tag{3.17}
\]

where $R_0 = \beta/\gamma$ is the reproduction number. To model uncertainty in the repro-
duction number, \( R_0 \) was written in the form

\[
R_0 = r_0 + p\theta, \tag{3.18}
\]

where \( \theta \) is a random variable with a known probability distribution (similar to \( \xi \) in previous sections) and \( r_0 \) and \( p \) are (positive) constants. Substituting Equation (3.18) into Equations (3.16) and (3.17) gives

\[
\frac{dS}{d\tau} = -(r_0 + p\theta)SI \tag{3.19}
\]
\[
\frac{dI}{d\tau} = (r_0 + p\theta)SI - I. \tag{3.20}
\]

The stochastic Galerkin method was then applied to Equations (3.19) and (3.20), similar to Section 3.3.

It is not immediately obvious, but the way that Roberts has derived his equations (Equations (3.19) and (3.20)) actually makes time (\( \tau \)) rely on a random variable. In order to demonstrate how this has happened, once again consider the deterministic SIR model (Equations (3.14) and (3.15)). If there is uncertainty in \( R_0 \), this uncertainty could either come from the transmission rate (\( \beta \)), the recovery rate (\( \gamma \)) or both. To represent this uncertainty, \( \beta \) and \( \gamma \) are written as functions of random variables giving

\[
\beta = \beta_0 + a\xi_1 \tag{3.21}
\]
\[
\gamma = \gamma_0 + b\xi_2, \tag{3.22}
\]

where \( \beta_0, \gamma_0, a \) and \( b \) are (positive) constants and \( \xi_1 \) and \( \xi_2 \) are independent random variables with known probability distributions.

Substituting Equations (3.21) and (3.22) into Equations (3.14) and (3.15) gives

\[
\frac{dS}{dt} = -(\beta_0 + a\xi_1)SI
\]
\[
\frac{dI}{dt} = (\beta_0 + a\xi_1)SI - (\gamma_0 + b\xi_2)I
\]

and dividing through both equations by \( \gamma_0 + b\xi_2 \) gives

\[
\frac{1}{\gamma_0 + b\xi_2} \frac{dS}{dt} = -\frac{\beta_0 + a\xi_1}{\gamma_0 + b\xi_2} SI \tag{3.23}
\]
\[
\frac{1}{\gamma_0 + b\xi_2} \frac{dI}{dt} = \frac{\beta_0 + a\xi_1}{\gamma_0 + b\xi_2} SI - I. \tag{3.24}
\]

Comparing Equations (3.23) and (3.24) with Roberts’ equations (Equations (3.19))
and (3.20)) and assuming that

\[ \frac{\beta_0 + a\xi_1}{\gamma_0 + b\xi_2} = r_0 + p\theta \]  

(3.25)

then

\[ \tau = (\gamma_0 + b\xi_2)t, \]

which means that the time scale depends upon a random variable. Additionally, the random variable associated with time, \( \xi_2 \), is not the random variable associated with \( R_0, \theta \). However, \( \theta \) depends upon \( \xi_2 \), as shown in Equation (3.25). This would make the new time scale (\( \tau \)) very difficult to revert back to the normal time scale (\( t \)).

If it is assumed that the uncertainty in \( R_0 \) only comes from \( \beta \) (\( b = 0 \)), Equations (3.23) and (3.24) simplify to Roberts’ equations by letting

\[ r_0 + p\theta = \frac{\beta_0}{\gamma_0} + \frac{a}{\gamma_0}\xi_1. \]

This also means that there is no random variable introduced into the time scale. However, if the uncertainty is only associated with \( \beta \), the problem is exactly the same as described earlier in Section 3.2.

It is important to note that the probability distribution of \( R_0 \) can also take complicated forms, even if the distributions for \( \beta \) and \( \gamma \) are simple. For example, consider that uncertainty has been included in \( \beta \) and \( \gamma \) such that

\[ \beta = \beta_0 + a\xi_3 \]

\[ \gamma = \gamma_0 + b\xi_4, \]

where \( \xi_3 \) and \( \xi_4 \) are uniformly distributed random variables on \([-1, 1]\). The probability distributions of \( \beta \) and \( \gamma \) can be seen in Figure 3.6, panels (a) and (b), when \( \beta_0 = 6, \gamma_0 = 3, a = 3 \) and \( b = 1 \). The probability density function for \( R_0 \), \( P(R_0) \), can be determined analytically and is given by

\[
P(R_0) = \begin{cases} 
  \frac{16R_0^2 - 9}{24R_0^2} & \text{if } 0.75 \leq z < 1.5 \\
  0.5 & \text{if } 1.5 \leq z < 2.25 \\
  \frac{81 - 4R_0^2}{24R_0^2} & \text{if } 2.25 \leq z \leq 4.5 \\
  0 & \text{otherwise.}
\end{cases}
\]

Despite simple uniform distributions for \( \beta \) and \( \gamma \), the probability distribution for \( R_0 \) is not a uniform distribution and is more complicated. The probability distribution for \( R_0 \) can be seen in Figure 3.6(c). Therefore, if \( R_0 \) has a uniform distribution, as
considered in Roberts’ paper, the distributions of $\beta$ and $\gamma$ would not be uniform and would most likely be quite complicated distributions.

Rather than trying to introduce uncertainty directly into $R_0$, it would be better to simply consider uncertainty in both $\beta$ and $\gamma$. While applying the stochastic Galerkin method will be slightly more difficult due to multiple random variables, it allows for much greater flexibility in the uncertainty distributions of $\beta$ and $\gamma$. This will then allow greater flexibility in the distribution of $R_0$. The variance of the predictions obtained from the SIR model can also be deconstructed to determine the proportion of the variance due to each of the uncertain parameters. This is done using Sobol indices and is investigated in Section 3.7. If uncertainty is only considered in $R_0$, the variance cannot be deconstructed into its parts.
An $SIR$ model with uncertainty in $\beta$ and $\gamma$ is presented by Chen-Charpentier and Stanescu [13] as well as being investigated in Section 3.6. However, neither Roberts nor Chen-Charpentier and Stanescu consider uncertainty in the initial conditions of their epidemic models. The effect of uncertainty in the initial conditions of the susceptible and infected populations is investigated in Section 3.5.

### 3.5 $SIR$ model with uncertainty in the initial conditions

In Section 3.3, an $SIR$ with uncertainty in one of its parameters was investigated and solved using the stochastic Galerkin method whereas in this section, an $SIR$ model with uncertainty in the initial conditions will be investigated. To represent the uncertainty in the initial conditions, the initial conditions will now be written as functions of random variables. Therefore, the initial conditions are now given by

\[
S_{\text{initial}} = S(0, \xi) = 1 - C - g\xi \\
I_{\text{initial}} = I(0, \xi) = C + g\xi \\
R_{\text{initial}} = R(0, \xi) = 0,
\]

where $0 < C \leq 1$, $g$ is a positive scaling parameter and $\xi$ is a random variable with a known probability distribution.

It is important that the values of $C$ and $g$, as well as the probability distribution for $\xi$, are chosen carefully so that for all possible values of $\xi$, $0 < I_{\text{initial}} \leq 1$. This is required so that the $SIR$ model returns sensible predictions. For example, a gamma distributed random variable should not be used to represent uncertainty in the initial conditions as a gamma distributed random variable can take any non-negative value (ie any value on the interval $[0, \infty)$). Hence, even if a very small value of $g$ is used, it would still be possible for $I_{\text{initial}} > 1$. Therefore, in this example, $\xi$ will be chosen to be a uniform random variable on $[-1, 1]$ which has a probability density function $w(\xi) = 1/2$. While gamma distributed random variables cannot be used in this example, they will be used later in this chapter to represent uncertainty.

As in Section 3.3, the solutions for $S$ and $I$ are written in the form

\[
S(t, \xi) = \sum_{i=0}^{\infty} S_i(t)\Psi_i(\xi) \approx \sum_{i=0}^{P} S_i(t)\Psi_i(\xi) \\
I(t, \xi) = \sum_{i=0}^{\infty} I_i(t)\Psi_i(\xi) \approx \sum_{i=0}^{P} I_i(t)\Psi_i(\xi)
\]

and truncated at the $P$th order where $\Psi_i(\xi)$ are the Legendre polynomials. Substi-
tuting into the SIR model (Equations (2.12) and (2.13)) gives

\[
\sum_{i=0}^{P} \frac{dS_i(t)}{dt}\Psi_i(\xi) = -\beta \sum_{i=0}^{P} S_i(t)\Psi_i(\xi) \sum_{j=0}^{P} I_j(t)\Psi_j(\xi)
\]
\[
+ \sum_{i=0}^{P} \frac{dI_i(t)}{dt}\Psi_i(\xi) = \beta \sum_{i=0}^{P} S_i(t)\Psi_i(\xi) \sum_{j=0}^{P} I_j(t)\Psi_j(\xi)
\]
\[- \gamma \sum_{i=0}^{P} I_i(t)\Psi_i(\xi).
\]

Multiplying by \(\Psi_k(\xi)\) and integrating over the probability space gives

\[
\sum_{i=0}^{P} \frac{dS_i(t)}{dt} \langle \Psi_i(\xi), \Psi_k(\xi) \rangle = -\beta \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t)I_j(t) \langle \Psi_i(\xi)\Psi_j(\xi), \Psi_k(\xi) \rangle
\]
\[
+ \sum_{i=0}^{P} \frac{dI_i(t)}{dt} \langle \Psi_i(\xi), \Psi_k(\xi) \rangle = \beta \sum_{i=0}^{P} \sum_{j=0}^{P} S_i(t)I_j(t) \langle \Psi_i(\xi)\Psi_j(\xi), \Psi_k(\xi) \rangle
\]
\[- \gamma \sum_{i=0}^{P} I_i(t) \langle \Psi_i(\xi), \Psi_k(\xi) \rangle \]  
(3.26)

where

\[\langle \Psi_i(\xi), \Psi_k(\xi) \rangle = \frac{1}{2} \int_{-1}^{1} \Psi_i(\xi)\Psi_k(\xi)d\xi.\]

By evaluating the inner products, a deterministic system of \(2(P + 1)\) ordinary differential equations can be determined. However, before this system of equations can be solved, initial conditions for \(S_i(t)\) and \(I_i(t)\) need to be determined. The initial condition for the fraction of infected individuals is given by

\[I(0, \xi) = C + g\xi = \sum_{i=0}^{P} I_i(0)\Psi_i(\xi).\]

Multiplying through by \(\Psi_k(\xi)\) and integrating over the probability space gives

\[C \int_{-1}^{1} \Psi_k(\xi)\frac{1}{2}d\xi + g \int_{-1}^{1} \xi\Psi_k(\xi)\frac{1}{2}d\xi = \sum_{i=0}^{P} I_i(0) \int_{-1}^{1} \Psi_i(\xi)\Psi_k(\xi)\frac{1}{2}d\xi.\]

However, as

\[\int_{-1}^{1} \Psi_k(\xi)\frac{1}{2}d\xi = \begin{cases} 
1 & \text{for } k = 0 \\
0 & \text{for } k \geq 1
\end{cases} \]

and

\[\int_{-1}^{1} \xi\Psi_k(\xi)\frac{1}{2}d\xi = \begin{cases} 
\frac{1}{3} & \text{for } k = 1 \\
0 & \text{otherwise}
\end{cases} \]
the initial conditions for $I_i(t)$ are simply given by

$$I_i(0) = \begin{cases} 
C & \text{if } i = 0 \\
g & \text{if } i = 1 \\
0 & \text{if } i > 1.
\end{cases}$$

Using a similar approach, the initial conditions for $S_i(t)$ are given by

$$S_i(0) = \begin{cases} 
1 - C & \text{if } i = 0 \\
-g & \text{if } i = 1 \\
0 & \text{if } i > 1.
\end{cases}$$

Figure 3.7 shows a comparison of the solutions obtained with Monte Carlo sampling (with $10^5$ trials) and the stochastic Galerkin method, with the stochastic Galerkin solution being truncated at the first, third and fifth order expansions ($P = 1$, $P = 3$ and $P = 5$, respectively). The probability distributions for $S_{\text{initial}}$ and $I_{\text{initial}}$ can be seen in Figure 3.8. Even using only a first order expansion, the mean solution is very close to the Monte Carlo solution. However, to obtain an accurate solution for the variance, higher order expansions are needed. Using a fifth order expansion ($P = 5$), the variance calculated using the Monte Carlo method and the stochastic Galerkin method are almost identical.

It is interesting to note that at the start of the epidemic ($t = 0$), there is already variance in the mean due to the uncertainty in the initial conditions. The peak in the variance also occurs before the peak of the epidemic, whereas in Figure 3.5 (where there was only uncertainty in $\beta$), the peak in the variance was at approximately the same time as the peak in the epidemic. By the time the epidemic reaches its peak, the variance has significantly decreased from its peak. Also, as noted previously, there are two peaks in the variance of the infected population. After reaching its first peak, the variance decreases to almost zero, which is then followed by a much smaller peak, before again returning to almost zero.

\section*{3.6 $SIR$ model with uncertainty in two parameters}

In previous sections, the stochastic Galerkin method was investigated when there was one uncertain parameter or uncertainty in one of the initial conditions of an $SIR$ model. Therefore, only one random variable was needed to represent the uncertainty. In this section, the stochastic Galerkin method will be investigated when there are two uncertain parameters in the $SIR$ model.

As before, to represent the uncertainty in the parameters, the parameters will
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.7: Comparison of stochastic Galerkin solution with Monte Carlo sampling for SIR model with uncertainty in the initial conditions. Blue is susceptible and red is infected. Solid lines are Monte Carlo solutions and dashed lines are stochastic Galerkin solutions. $\beta = 4$, $\gamma = 1$, $\xi$ is a uniform distribution on $[-1, 1]$, $I_{initial} = 0.1 + 0.08\xi$. 
be written as functions of random variables. Therefore $\beta$ and $\gamma$ are now written in the form

$$
\beta = \beta_0 + a\xi_1 \\
\gamma = \gamma_0 + b\xi_2,
$$

where $\xi_1$ and $\xi_2$ are independent random variables with probability density functions, $w_1(\xi_1)$ and $w_2(\xi_2)$, respectively, and sample spaces $\Omega_1$ and $\Omega_2$, respectively. For now, the stochastic Galerkin method will be applied without considering specific distributions for the random variables, $\xi_1$ and $\xi_2$. In later sections, random variables with uniform, gamma and normal distributions are considered.

In order to implement the stochastic Galerkin method, the solutions for $S$ and $I$ are now written in the form [25]

$$
S(t, \xi_1, \xi_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} S_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2) \approx \sum_{i=0}^{P} \sum_{j=0}^{P-i} S_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2)
$$

$$
I(t, \xi_1, \xi_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} I_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2) \approx \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2),
$$

where $\Psi_i(\xi_1)$ and $\Phi_j(\xi_2)$ are appropriately chosen orthogonal polynomials. Note that $S$ and $I$ are now functions of not only time, but both random variables ($\xi_1$ and $\xi_2$). As before, for practicality, the expansions have been truncated at the $P$th order. However, now that there are two random variables, a $P$th order expansion now has $\binom{P+2}{2}$ terms. For example, a second order expansion ($P = 2$) now results in 12 deterministic ODEs (6 for $S$ and 6 for $I$) while a third order expansion ($P = 3$) now results in 20 deterministic ODEs (10 for $S$ and 10 for $I$). A full list of terms up to third order can be seen in Table 3.1. By summing the subscript numbers of a particular term, it can quickly be determined what order term it is. For example, $S_{12}(t)$ is a third order term while $I_{31}(t)$ is a fourth order term.

Substituting the form of the solution (Equations (3.28)) and the two parameters
Chapter 3. Applying the stochastic Galerkin method to the SIR model

<table>
<thead>
<tr>
<th>Order</th>
<th>Associated Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>zero</td>
<td>$S_{00}(t)$, $I_{00}(t)$</td>
</tr>
<tr>
<td>first</td>
<td>$S_{10}(t)$, $I_{10}(t)$, $S_{01}(t)$, $I_{01}(t)$</td>
</tr>
<tr>
<td>second</td>
<td>$S_{20}(t)$, $I_{20}(t)$, $S_{11}(t)$, $I_{11}(t)$, $S_{02}(t)$, $I_{02}(t)$</td>
</tr>
<tr>
<td>third</td>
<td>$S_{30}(t)$, $I_{30}(t)$, $S_{21}(t)$, $I_{21}(t)$, $S_{12}(t)$, $I_{12}(t)$, $S_{03}(t)$, $I_{03}(t)$</td>
</tr>
</tbody>
</table>

Table 3.1: Terms involved in different order expansions when two parameters contain uncertainty in an SIR model.

Applying a Galerkin projection by multiplying by $\Psi_u(\xi_1)\Phi_v(\xi_2)$ ($u = 0, 1, 2, \ldots, P$ and $v = 0, 1, 2, \ldots, (P - u)$) and integrating over the probability space gives

$$\sum_{i=0}^{P} \sum_{j=0}^{P-i} \frac{dS_{ij}(t)}{dt} \Psi_i(\xi_1)\Phi_j(\xi_2)$$

$$= - (\beta_0 + a\xi_1) \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij}(t)I_{mn}(t)\Psi_i(\xi_1)\Phi_j(\xi_2)\Psi_m(\xi_1)\Phi_n(\xi_2)$$

$$\sum_{i=0}^{P} \sum_{j=0}^{P-i} \frac{dI_{ij}(t)}{dt} \Psi_i(\xi_1)\Phi_j(\xi_2)$$

$$= (\beta_0 + a\xi_1) \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij}(t)I_{mn}(t)\Psi_i(\xi_1)\Phi_j(\xi_2)\Psi_m(\xi_1)\Phi_n(\xi_2)$$

$$- (\gamma_0 + b\xi_2) \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2).$$

Applying a Galerkin projection by multiplying by $\Psi_u(\xi_1)\Phi_v(\xi_2)$ ($u = 0, 1, 2, \ldots, P$ and $v = 0, 1, 2, \ldots, (P - u)$) and integrating over the probability space gives

$$\sum_{i=0}^{P} \sum_{j=0}^{P-i} \frac{dS_{ij}(t)}{dt} \langle \Psi_i\Phi_j, \Psi_u\Phi_v \rangle = - \beta_0 \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij}I_{mn} \langle \Psi_i\Phi_j, \Psi_m\Phi_n, \Psi_u\Phi_v \rangle$$

$$- a \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij}I_{mn} \langle \xi_1\Psi_i\Phi_j, \Psi_m\Phi_n, \Psi_u\Phi_v \rangle$$
Chapter 3. Applying the stochastic Galerkin method to the SIR model

where the inner product \( \langle F, G \rangle \) is given by

\[
\langle F(\xi_1, \xi_2), G(\xi_1, \xi_2) \rangle = \int_{\Omega_2} \int_{\Omega_1} F(\xi_1, \xi_2) G(\xi_1, \xi_2) w(\xi_1) w(\xi_2) d\xi_1 d\xi_2.
\]

For simplicity and readability, the inner product \( \langle F, G \rangle \) has been expressed as a double integral. However, when calculating the inner product, as each orthogonal polynomial is only a function of one of the random variables, the inner product can be calculated using the product of two single integrals rather than a double integral. For example,

\[
\langle \Psi_i \Phi_j, \Psi_u \Phi_v \rangle = \int_{\Omega_2} \int_{\Omega_1} \Psi_i(\xi_1) \Phi_j(\xi_2) \Psi_u(\xi_1) \Phi_v(\xi_2) w(\xi_1) w(\xi_2) d\xi_1 d\xi_2
\]

\[
= \left[ \int_{\Omega_1} \Psi_i(\xi_1) \Psi_u(\xi_1) w(\xi_1) d\xi_1 \right] \left[ \int_{\Omega_2} \Phi_j(\xi_2) \Phi_v(\xi_2) w(\xi_2) d\xi_2 \right].
\]

Additionally, as \( \Psi_i(\xi_1) \) and \( \Phi_j(\xi_2) \) are orthogonal polynomials, many of the integrals, and therefore many of the inner products, trivially evaluate to zero. This gives

\[
\frac{dS_{uv}(t)}{dt} = -\frac{\beta_0}{K_{uv}} \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij} I_{mn} \langle \Psi_i \Phi_j, \Psi_m \Phi_n, \Psi_u \Phi_v \rangle
\]

\[
\frac{dI_{uv}(t)}{dt} = \frac{\beta_0}{K_{uv}} \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij} I_{mn} \langle \Psi_i \Phi_j, \Psi_m \Phi_n, \Psi_u \Phi_v \rangle + \frac{a}{K_{uv}} \sum_{i=0}^{P} \sum_{j=0}^{P-i} \sum_{m=0}^{P} \sum_{n=0}^{P-m} S_{ij} I_{mn} \langle \xi_i \Psi_j \Phi_m \Phi_n, \Psi_u \Phi_v \rangle
\]

\[
- \gamma_0 I_{uv} - \frac{b}{K_{uv}} \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij} \langle \xi_2 \Psi_i \Phi_j, \Psi_u \Phi_v \rangle.
\]
where
\[ K_{uv} = \int_{\Omega_2} \int_{\Omega_1} (\Psi_u(\xi_1))^2 (\Phi_v(\xi_2))^2 w_1(\xi_1) w_2(\xi_2) d\xi_1 d\xi_2 \]
\[ = \left[ \int_{\Omega_1} (\Psi_u(\xi_1))^2 w_1(\xi_1) d\xi_1 \right] \left[ \int_{\Omega_2} (\Phi_v(\xi_2))^2 w_2(\xi_2) d\xi_2 \right]. \]

This gives a system of \( 2 \binom{P+2}{2} \) deterministic differential equations that can easily be solved using a numerical ODE solver. While uncertainty was introduced to the SIR model by using two independent random variables, the final system of equations is deterministic and only needs to be solved once.

A MATLAB routine that symbolically generates the \( 2 \binom{P+2}{2} \) deterministic differential equations for an SIR model with uncertainty in \( \beta \) and \( \gamma \) can be seen in Appendix A. While for low order expansions, it is not necessary to generate the deterministic ODEs symbolically due to the small computational cost, it is advantageous for higher order expansions. Once the deterministic ODEs have been generated, the scaling parameters can easily be changed and updated predictions can be calculated without having to reapply the stochastic Galerkin method or rederive the deterministic ODEs.

Assuming that there is no uncertainty in the initial conditions, and that there is initially no one in the removed compartment, the initial conditions are given by
\[ S(0, \xi_1, \xi_2) = 1 - I_{\text{initial}} \]
\[ I(0, \xi_1, \xi_2) = I_{\text{initial}}. \]

From these initial conditions, the initial conditions for the stochastic Galerkin solution can be determined. Substituting the initial condition for the fraction of infected individuals into the stochastic Galerkin expansion gives
\[ I(0, \xi_1, \xi_2) = I_{\text{initial}} = \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij}(0) \Psi_i(\xi_1) \Phi_j(\xi_2). \]

Performing a Galerkin projection on the initial condition gives
\[ I_{\text{initial}} \int_{\Omega_2} \int_{\Omega_1} \Psi_u(\xi_1) \Phi_v(\xi_2) w_1 w_2 d\xi_1 d\xi_2 = \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij}(0) \int_{\Omega_2} \int_{\Omega_1} \Psi_i(\xi_1) \Phi_j(\xi_2) \Phi_u(\xi_1) \Psi_v(\xi_2) w_1 w_2 d\xi_1 d\xi_2. \]

However, as
\[ \int_{\Omega_2} \int_{\Omega_1} \Psi_u(\xi_1) \Phi_v(\xi_2) w_1 w_2 d\xi_1 d\xi_2 = \begin{cases} 1 & \text{if } u = 0 \text{ and } v = 0 \\ 0 & \text{otherwise} \end{cases} \]
the initial conditions for $I_{ij}(t)$ are simply given by

$$I_{ij}(0) = \begin{cases} I_{\text{initial}} & \text{if } i = 0 \text{ and } j = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the initial conditions for $S_{ij}(t)$ are given by

$$S_{ij}(0) = \begin{cases} 1 - I_{\text{initial}} & \text{if } i = 0 \text{ and } j = 0 \\ 0 & \text{otherwise.} \end{cases}$$

As in the previous sections, the mean and variance can be determined directly from the stochastic Galerkin solution. For example, the mean fraction of infected individuals $E[I(t, \xi_1, \xi_2)]$ is given by

$$E[I(t, \xi_1, \xi_2)] = E \left[ \sum_{i=0}^{P} \sum_{j=0}^{P} I_{ij}(t) \Psi_i(\xi_1) \Phi_j(\xi_2) \right]$$

$$= \sum_{i=0}^{P} \sum_{j=0}^{P} I_{ij}(t) \int_{\Omega_2} \int_{\Omega_1} \Psi_i(\xi_1) \Phi_j(\xi_2) w_1(\xi_1) w_2(\xi_2) d\xi_1 d\xi_2$$

$$= I_{00}(t)$$

and the variance ($\text{Var}[I(t, \xi_1, \xi_2)]$) is given by

$$\text{Var}[I(t, \xi_1, \xi_2)] = E[I(t, \xi_1, \xi_2)^2] - (E[I(t, \xi_1, \xi_2)])^2$$

$$= E \left[ \sum_{i=0}^{P} \sum_{j=0}^{P} \sum_{m=0}^{P} \sum_{n=0}^{P} I_{ij}(t) I_{mn}(t) \Psi_i(\xi_1) \Phi_j(\xi_2) \Psi_m(\xi_1) \Phi_n(\xi_2) \right] - (I_{00}(t))^2$$

$$= \sum_{i=0}^{P} \sum_{j=0}^{P} (I_{ij}(t))^2 \langle (\Psi_i(\xi_1))^2, (\Phi_j(\xi_2))^2 \rangle - (I_{00}(t))^2.$$
Therefore the mean solution is simply given by the zero order term \((I_{00}(t))\) while the variance can be quickly calculated from the sum of the squares of the non-zero order terms.

By incorporating uncertainty into \(\beta\) and \(\gamma\), it has allowed for much greater flexibility in the representation of the uncertainty compared with only considering uncertainty in the basic reproduction number, \(R_0\), as considered by Roberts [12]. Additionally, as seen in Section 3.4, even if \(\beta\) and \(\gamma\) have simple uniform distributions, the distribution for \(R_0\) can be much more complicated. By introducing scaling parameters \((\beta_0, \gamma_0, a\) and \(b\)), the final system of deterministic ODEs, from the stochastic Galerkin method, is more general and can be used for different probability distributions of the same type. This is a significant advantage over the approach used by Chen-Charpentier and Stanescu [13], which would need additional calculations if the probability distributions of the uncertain parameters were to change, as will be discussed further in later sections of this chapter.

### 3.7 Determination of Sobol indices

Now that there are multiple sources of uncertainty, it can be useful to know what fraction of the variance is due to the each of the sources of uncertainty. If the majority of the variance is due to a single source of uncertainty, additional time and effort could be devoted to determining the value of the parameter, or at the very least, narrowing the relative size of the random variable representing the uncertainty. Alternatively, if the uncertainty in a particular parameter is shown to be responsible for only a small fraction of the overall variance, there is no need to spend time or resources on more accurately measuring the parameter. Also, if the uncertainty in the parameter is responsible for a small fraction of the variance, the parameter could simply be considered a constant, as it has very little impact on the final results. By reducing the number of parameters with uncertainty, the stochastic Galerkin solution becomes much easier to calculate.

Sobol indices can be used to decompose the variance caused by the uncertainty in the different parameters. The Sobol indices can be calculated through Monte Carlo sampling, but can also be determined directly from the stochastic Galerkin expansion with very little additional calculation [72].

For example, the stochastic Galerkin expansion for the fraction of infected individuals with two parameters with uncertainty is written in the form

\[
I(t, \xi_1, \xi_2) = \sum_{i=0}^{P} \sum_{j=0}^{P-i} I_{ij}(t)\Psi_i(\xi_1)\Phi_j(\xi_2)
\]
Chapter 3. Applying the stochastic Galerkin method to the SIR model

and its variance (Var) is usually written in the form

\[ \text{Var}[I(t, \xi_1, \xi_2)] = \sum_{i=0}^{P} \sum_{j=0}^{P-i} (I_{ij}(t))^2 \langle (\Psi_i(\xi_1))^2, (\Phi_j(\xi_2))^2 \rangle - (I_{00}(t))^2. \]

However the variance can also be rewritten in the form

\[ \text{Var}[I(t, \xi_1, \xi_2)] = \sum_{i=1}^{P} (I_{i0}(t))^2 \langle \Psi_i(\xi_1), \Psi_i(\xi_1) \rangle 
+ \sum_{j=1}^{P} (I_{0j}(t))^2 \langle \Phi_j(\xi_2), \Phi_j(\xi_2) \rangle 
+ \sum_{i=1}^{P} \sum_{j=1}^{P-i} (I_{ij}(t))^2 \langle (\Psi_i(\xi_1))^2, (\Phi_j(\xi_2))^2 \rangle. \]

The variance has now been decomposed into three distinct components. As the first component contains only the random variable \( \xi_1 \), this is the variance due to the uncertainty in \( \beta \). The second component contains only the random variable \( \xi_2 \) and is therefore the variance due to the uncertainty in \( \gamma \). The third component contains both \( \xi_1 \) and \( \xi_2 \) and is therefore the variance caused by the interaction of the uncertainty in \( \beta \) and \( \gamma \).

By dividing each component by the total variance, the Sobol indices for the infected population can be determined and are given by

\[
S^I_\beta = \frac{\sum_{i=1}^{P} (I_{i0}(t))^2 \langle \Psi_i(\xi_1), \Psi_i(\xi_1) \rangle}{\text{Var}[I(t, \xi_1, \xi_2)]},
\]

\[
S^I_\gamma = \frac{\sum_{j=1}^{P} (I_{0j}(t))^2 \langle \Phi_j(\xi_2), \Phi_j(\xi_2) \rangle}{\text{Var}[I(t, \xi_1, \xi_2)]},
\]

\[
S^I_{\beta\gamma} = \frac{\sum_{i=1}^{P} \sum_{j=1}^{P-i} (I_{ij}(t))^2 \langle (\Psi_i(\xi_1))^2, (\Phi_j(\xi_2))^2 \rangle}{\text{Var}[I(t, \xi_1, \xi_2)]},
\]

where \( S^I_\beta \) is the fraction of the variance caused by the uncertainty in \( \beta \), \( S^I_\gamma \) is the fraction of the variance caused by the uncertainty in \( \gamma \) and \( S^I_{\beta\gamma} \) is the fraction of the variance caused by the interaction of the uncertainty in \( \beta \) and \( \gamma \). As each Sobol index represents a fraction of the total variance, \( S^I_\beta + S^I_\gamma + S^I_{\beta\gamma} = 1 \). Similar expressions can be determined for the Sobol indices of the susceptible population, \( S^S_\beta, S^S_\gamma \) and \( S^S_{\beta\gamma} \).

### 3.8 Comparison of results

In Section 3.6, the stochastic Galerkin method was applied to the SIR with uncertainty in both \( \beta \) and \( \gamma \) and in Section 3.7, the Sobol indices for the SIR model
were calculated. Throughout these two sections, no particular probability distribution was considered for the random variables. By doing this, the equations derived would work for any probability distribution.

In this section, uniform, gamma and normal distributions will be considered for the random variables in the SIR model. In each case, the stochastic Galerkin solution will be compared to the solution obtained from Monte Carlo sampling. Additionally, the Sobol indices will be calculated. This was not necessary for the SIR model in Section 3.2, as there was uncertainty only in $\beta$ and therefore all the variance was due to this uncertainty. Now that there are two sources of uncertainty, it is helpful to calculate the Sobol indices to help understand which source of uncertainty is responsible for the variance.

### 3.8.1 Comparison of results: uniform distributions

In this section, it will be assumed that $\xi_1$ and $\xi_2$ are independent random variables with uniform distributions on $[-1,1]$. As $w_1(\xi_1) = 1/2$ and $w_2(\xi_2) = 1/2$, the Legendre polynomials will be used as basis functions for the stochastic Galerkin solution.

It is important to note that while the random variables, $\xi_1$ and $\xi_2$, are uniformly distributed on $[-1,1]$, the parameters $\beta$ and $\gamma$ can have any uniform distribution by simply choosing appropriate values for $\beta_0$, $\gamma_0$, $a$ and $b$. This means that the final system of deterministic equations from the stochastic Galerkin method can be used for any uniform distributions for $\beta$ and $\gamma$. However, care must be taken to ensure that $\beta$ and $\gamma$ remain positive for all values of $\xi_1$ and $\xi_2$.

Figure 3.9 shows the probability distributions for $\beta$ and $\gamma$ when $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$ and $b = 0.5$. Figure 3.10 shows a comparison of results obtained from Monte Carlo sampling and the stochastic Galerkin method with different order expansions. As the Monte Carlo solution was obtained using $10^5$ trials, it will be considered to be very close to the ‘exact’ solution.

Even using only a first order expansion, the mean solution obtained from the stochastic Galerkin method is very close to the ‘exact’ solution. However, using such a low order expansion does not give a good approximation to the variance. As the order of the stochastic Galerkin expansion is increased to second order, the variance obtained from the stochastic Galerkin method is much closer to the exact solution. Using a fourth order expansion, the results obtained from the stochastic Galerkin method are almost identical to the ‘exact’ solution. It is important to note that while a fourth order expansion may seem very easy to use, it actually results in a final system of 30 ordinary differential equations as the number of equations is given by $2^{(P+2)/2}$. If the expansion were to be increased to fifth order, this would result in a final system of 42 differential equations.

Figure 3.11 shows the Sobol indices for both the susceptible population and the
infected population when the random variables are uniformly distributed. For the susceptible population, the variance is initially entirely due to the uncertainty in $\beta$ with the variance due to $\gamma$ or the interaction of $\beta$ and $\gamma$ being almost zero. As time passes, the variance due to the uncertainty in $\beta$ decreases while the variance due to $\gamma$ and the interaction of $\beta$ and $\gamma$ increases. By the end of the epidemic ($t = 6$), the variance due to $\beta$ is almost the same as the variance due to $\gamma$.

For the infected population, once again the variance at the beginning of the epidemic is mostly due to the uncertainty in $\beta$. As time passes, this rapidly changes and by $t = 2$, the variance from the uncertainty in $\beta$ is about the same as the variance due to $\gamma$. By the end of the epidemic ($t = 6$), the majority of the variance (approximately 70%) comes from the uncertainty in $\gamma$. It is also interesting to note that there is almost no contribution to the variance due to the interaction of the uncertainty in $\beta$ and $\gamma$. As the rate of infection has nothing to do with the recovery rate of the disease, it is not surprising that there is little interaction between the uncertainty in these variables which is confirmed by the Sobol indices.

### 3.8.2 Comparison of results: gamma distribution

In this section, it will be assumed that $\xi_1$ and $\xi_2$ are independent random variables that have gamma distributions with shape and scale parameters equal to one. Therefore, $w_1(\xi_1) = \exp(-\xi_1)$ and $w_2(\xi_2) = \exp(-\xi_2)$ and their support is $\xi_1 \in [0, \infty)$ and $\xi_2 \in [0, \infty)$. As the random variables are now gamma distributed, the Laguerre polynomials will be used as basis functions for the stochastic Galerkin expansion as the weight function of the Laguerre polynomials matches the probability density function of the random variables.
Figure 3.10: Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$, $b = 0.5$, $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{\text{initial}} = 0.01$. 

Chapter 3. Applying the stochastic Galerkin method to the SIR model
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.11: Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1.5$, $b = 0.5$, $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{\text{initial}} = 0.01$, $P = 4$. 
The first few Laguerre polynomials, \( L_i(\xi_1) \), are given by [62]

\[
\begin{align*}
L_0(\xi_1) &= 1 \\
L_1(\xi_1) &= -\xi_1 + 1 \\
L_2(\xi_1) &= ((\xi_1)^2 - 4\xi_1 + 2)/2 \\
L_3(\xi_1) &= (-3(\xi_1)^3 + 9(\xi_1)^2 - 18\xi_1 + 6)/6
\end{align*}
\]

and their recurrence relation is given by [62]

\[(i + 1)L_{i+1}(\xi_1) = (2i + 1 - \xi_1)L_i(\xi_1) - iL_{i-1}(\xi_1).\]

In the previous section where uniformly distributed random variables were used, the values of \( \beta_0, \gamma_0, a \) and \( b \) were chosen so that for all values of the random variables, \( R_0 > 1 \). However, as the random variables are now gamma distributed, it is no longer possible to choose values for the parameters such that \( R_0 > 1 \). This means that for some values of the random variables, the value of \( R_0 \) could be less than one, with the result that an epidemic does not occur. This could lead to misleading predictions from the SIR model.

Consider the case, using gamma distributed random variables (and appropriate values for \( \beta_0, \gamma_0, a \) and \( b \)), where it was found that roughly half the time \( R_0 \leq 1 \) and the other half of the time \( R_0 > 1 \). In this case, calculating the mean prediction would not be helpful. If, when the disease started to spread, the actual value of \( R_0 \leq 1 \), then no epidemic would occur and the mean prediction would severely overestimate the number infected. If instead, the actual value of \( R_0 > 1 \), the mean prediction would severely underestimate the epidemic. In either case, the mean prediction is not helpful. Therefore, when using gamma distributed random variables, \( \beta_0, \gamma_0, a \) and \( b \) should be carefully chosen such that there is only a small chance that \( R_0 < 1 \). This will ensure that the mean prediction is meaningful.

Figure 3.12 shows the probability distributions for \( \beta \) and \( \gamma \) when \( \beta_0 = 4, \gamma_0 = 1, a = 0.41 \) and \( b = 0.41 \). Figure 3.13 shows a comparison of results obtained from Monte Carlo sampling with \( 10^5 \) trials and the stochastic Galerkin method with different order expansions. As before, the Monte Carlo solution will be considered to be the ‘exact’ solution.

Even using only a first order stochastic Galerkin expansion, the mean solution is very close to the ‘exact’ solution and the variance is relatively close to the ‘exact’ solution. However, increasing the stochastic Galerkin expansion to second order actually gives a worse result for the variance when compared to the first order expansion. When the expansion is increased to third order, the variance obtained from the stochastic Galerkin method is very close to the ‘exact’ solution.

Figure 3.14 shows the Sobol indices, for both susceptible and infected populations, when the random variables are gamma distributed. For the susceptible
Chapter 3. Applying the stochastic Galerkin method to the SIR model

3.8.3 Comparison of results: normal distribution

In this section, it will be assumed that $\xi_1$ and $\xi_2$ are independent random variables that have normal distributions with zero mean and unit variance. Therefore, $w(\xi_1) = 1/\sqrt{2\pi} \exp(-\xi_1^2/2)$ and $w(\xi_2) = 1/\sqrt{2\pi} \exp(-\xi_2^2/2)$ and their support is $\xi_1 \in (-\infty, \infty)$ and $\xi_2 \in (-\infty, \infty)$.

It is important to note that while the random variables, $\xi_1$ and $\xi_2$, are normally distributed with zero mean and unit variance, the parameters $\beta$ and $\gamma$ can have any normal distribution by choosing suitable values for $\beta_0$, $\gamma_0$, $a$ and $b$. The mean and variance of $\beta$ are $\beta_0$ and $a$, respectively, and the mean and variance of $\gamma$ are $\gamma_0$ and $b$, respectively.

As the random variables are now normally distributed, the Hermite polynomials will be used as basis functions for the stochastic Galerkin expansion as the weight population, at the start of the epidemic, almost all of the variance is due to the uncertainty in $\beta$ with the uncertainty in $\gamma$ having almost no effect on the variance. By the end of the epidemic ($t = 6$), this has completely reversed with almost all of the variance being due to the uncertainty in $\gamma$. The interaction of uncertainty in $\beta$ and $\gamma$ has almost no effect on the variance.

For the infected population, at the start of the epidemic, the variance comes from both the uncertainty in $\beta$ and the uncertainty in $\gamma$, with both responsible for approximately half the variance. As the epidemic progresses, the variance due to the uncertainty in $\beta$ increases and then sharply decreases to almost zero at approximately $t = 2$. At the same time, the variance due to the uncertainty in $\gamma$ decreases after the start of the epidemic and then sharply increases. By $t = 2$, the uncertainty in $\gamma$ is almost entirely responsible for the variance. This decreases over time and by the end of the epidemic ($t = 6$), the uncertainty in $\gamma$ is responsible for approximately 60% of the variance.

\begin{figure}[h]
\centering
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{beta_distribution}
\caption{Probability distribution of $\beta$ for Figure 3.13. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$. $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one.}
\end{subfigure} \hspace{1cm}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{gamma_distribution}
\caption{Probability distribution of $\gamma$ for Figure 3.13. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$. $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one.}
\end{subfigure}
\end{figure}
Figure 3.13: Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$, $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one. $I_{initial} = 0.01$. 
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.14: Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.41$, $b = 0.41$, $\xi_1$ and $\xi_2$ are gamma distributed random variables with shape and scale parameters equal to one. $I_{\text{initial}} = 0.01$, $P = 3$. 
function of the Hermite polynomials matches the probability density functions of the random variables.

The first few Hermite polynomials, \(He_i(\xi_1)\), are given in Section 2.2.1 and their recurrence relation is given by [62]

\[
He_{i+1}(\xi_1) = \xi_1 He_i(\xi_1) - i He_{i-1}(\xi_1).
\]

While normal distributions are quite common, extreme care must be taken when using normally distributed random variables to represent the uncertainty in \(\beta\) and \(\gamma\). As normally distributed random variables can take any real number, no values of \(\beta_0, \gamma_0, a\) and \(b\) (other than \(a = b = 0\)) will ensure that \(\beta\) and \(\gamma\) remain positive for all values of the random variables. Therefore, while it cannot be guaranteed that \(\beta > 0\) and \(\gamma > 0\) for all values of the random variables, it should at least be the case that the probability of either parameter being negative is extremely small.

Figure 3.15 shows the probability distributions for \(\beta\) and \(\gamma\) when \(\beta_0 = 4, \gamma_0 = 1, a = 0.65\) and \(b = 0.2\). From the graphs, it is easy to see that it is very unlikely for \(\beta\) or \(\gamma\) to be negative. The probability of \(\beta\) or \(\gamma\) being negative, using the above values, can be calculated analytically. The probability that \(\beta\) is negative is

\[
\int_{-\infty}^{0} \frac{1}{a\sqrt{2\pi}} \exp\left(\frac{-(\xi_1 - \beta_0)^2}{2a^2}\right) d\xi_1 \approx 3.78 \times 10^{-10}
\]

and similarly the probability that \(\gamma\) is negative is

\[
\int_{-\infty}^{0} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-(\xi_2 - \gamma_0)^2}{2b^2}\right) d\xi_2 \approx 2.87 \times 10^{-7}.
\]

Therefore, while it is possible for \(\beta\) or \(\gamma\) to take negative values, due to using normally distributed random variables, the probability of either parameter being negative is extremely small.

Figure 3.16 shows a comparison of results obtained from Monte Carlo sampling with \(10^5\) trials and the stochastic Galerkin method with different order expansions. The parameter values used were \(\beta_0 = 4, \gamma_0 = 1, a = 0.65\) and \(b = 0.2\). Due to the large number of Monte Carlo trials, the Monte Carlo solution will be considered to be very close to the ‘exact’ solution. It is interesting to note that none of the \(10^5\) Monte Carlo trials resulted in a negative \(\beta\) or \(\gamma\) value.

It can also be seen from Figure 3.16 that using only a first order expansion, the mean solution from the stochastic Galerkin method is very close to the ‘exact’ solution. However, the variance obtained from the stochastic Galerkin method using a first order expansion only matches the variance obtained from Monte Carlo sampling up until the peak in the variance. The tails of the variance from the two different methods are quite different. Increasing the order of the stochastic Galerkin expansion to a second order, and then to a fourth order expansion gives significantly
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.15: Probability distributions for $\beta$ and $\gamma$ for Figure 3.16. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$. $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one, respectively.

better results for the variance. However, even when using a fourth order stochastic Galerkin expansion, the end of the tail of the variance ($5 \leq t \leq 6$) still slightly differs with that obtained from Monte Carlo sampling.

Figure 3.17 shows the Sobol indices for both susceptible and infected populations when the random variables are normally distributed. For the susceptible population, at the start of the epidemic, nearly all of the variance is due to the uncertainty in $\beta$. As the epidemic progresses, the Sobol index associated with $\beta$ declines while the Sobol index associated with $\gamma$ increases as well as the Sobol index associated with the interaction between $\beta$ and $\gamma$. By the end of the epidemic ($t = 6$), approximately 70% of the uncertainty is due to the uncertainty in $\beta$, 15% is due to the uncertainty in $\gamma$ and approximately 15% is due to the interaction between $\beta$ and $\gamma$ (with the Sobol index associated with the interaction between $\beta$ and $\gamma$ slightly higher than the Sobol index associated with $\gamma$). It is surprising that even at the end of the epidemic ($t = 6$), the majority of the variance is still due to the uncertainty in $\beta$. When the random variables were uniformly distributed, by the end of the epidemic, the uncertainty in $\beta$ only accounted for approximately 50% of the variance. In the case where the random variables were gamma distributed, by the end of the epidemic, the variance was almost entirely due to the uncertainty in $\gamma$.

For the infected population, at the start of the epidemic, about 90% of the variance is due to the uncertainty in $\beta$. As the epidemic progress, the Sobol index associated with $\beta$ starts to decline and reaches its minimum at approximately $t = 2.5$ before increasing again. By the end of the epidemic, it is at approximately 0.5. The Sobol index associated with $\gamma$ does the mirror opposite to the Sobol index associated with $\beta$. It first increases, reaches its maximum at approximately $t = 2.5$, before decreasing. At the end of the epidemic, it is at approximately 0.5. It is interesting to note that for the majority of the epidemic, there is almost no contribution to the variance due to the interaction between $\beta$ and $\gamma$, but there is a slight increase right at the end of the epidemic.
Figure 3.16: Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue are susceptible and red are infected. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$, $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one respectively. $I_{\text{initial}} = 0.01$. 

(a) Mean, $P = 1$  
(b) Variance, $P = 1$  
(c) Mean, $P = 2$  
(d) Variance, $P = 2$  
(e) Mean, $P = 4$  
(f) Variance, $P = 4$
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.17: Sobol indices for an SIR model with uncertainty in $\beta$ and $\gamma$. Blue and red are the Sobol indices associated with $\xi_1$ and $\xi_2$ respectively. Black is the Sobol index associated with the interaction of $\xi_1$ and $\xi_2$. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.65$, $b = 0.2$, $\xi_1$ and $\xi_2$ are normally distributed random variables with mean and variance equal to zero and one respectively. $I_{\text{initial}} = 0.01$, $P = 4$. 
3.8.4 Comparison of computation time vs accuracy

Throughout this chapter, it has been stated that the stochastic Galerkin method is faster than Monte Carlo sampling, as the final system of ODEs obtained through the stochastic Galerkin method only needs to be solved once. On the other hand, Monte Carlo sampling requires the model to be solved numerous times to obtain accurate results for the mean and the variance. While it is clear that solving a system of ODEs once is superior to the numerous solves needed for Monte Carlo sampling, no comparison of computational times has been given so far. In this section, the computational time and accuracy of both the stochastic Galerkin method and Monte Carlo sampling is investigated.

To compare the accuracy of each method, a Monte Carlo solution with $10^5$ trials will be considered as the ‘exact’ solution. While a Monte Carlo solution will never give the exact solution, with such a large number of trials, the Monte Carlo solution will be extremely close to the exact solution, as shown in Section 3.2.

For the stochastic Galerkin method, different order expansions were calculated and the time required to solve each order expansion was recorded. It was assumed that the final system of deterministic ODEs was already derived and did not add to the computational time of the stochastic Galerkin method. This is a reasonable assumption as the final system of deterministic ODEs only needs to be derived once. Additionally, as mentioned earlier, if the random variables are uniformly or normally distributed, the parameters $\beta$ and $\gamma$ can have any uniform or normal distribution by choosing appropriate values of $\beta_0$, $\gamma_0$, $a$ and $b$. For high order stochastic Galerkin expansions, deriving $2^{(P+2)}$ ODEs can become quite time consuming.

The stochastic Galerkin solutions were then compared to the ‘exact’ solution and the error calculated. The error was given by

$$\text{error} = \sum_{i=0}^{60} \left[ S^M_G(0.1i) - S^M_E(0.1i) \right]^2 + \left[ I^M_G(0.1i) - I^M_E(0.1i) \right]^2 + \left[ S^V_G(0.1i) - S^V_E(0.1i) \right]^2 + \left[ I^V_G(0.1i) - I^V_E(0.1i) \right]^2$$

where $S^M_G(t)$ and $I^M_G(t)$ are the stochastic Galerkin solutions for the mean susceptible and infected populations, respectively, at time $t$ and $S^M_E(t)$ and $I^M_E(t)$ are the ‘exact’ solutions for the mean susceptible and infected populations, respectively, at time $t$. Similarly, $S^V_G(t)$ and $I^V_G(t)$ are the variance in the susceptible and infected populations, respectively, calculated using the stochastic Galerkin solutions and $S^V_E(t)$ and $I^V_E(t)$ are the ‘exact’ variance in the susceptible and infected populations.

For the Monte Carlo method, the time required to run different numbers of trials was recorded, starting with 500 trials (which is very low for the Monte Carlo method) and increasing by 500 trials until 7500 trials was reached. These solutions were then compared to the ‘exact’ solutions (obtained using $10^5$ Monte Carlo trials) and the
error calculated using a similar error formula to Equation (3.30). Because the error from Monte Carlo sampling changes each time it is calculated due to having to ‘choose’ values for the random variables, the whole process was repeated 100 times so that the error at each number of trials could be averaged.

Figure 3.18 shows graphs of computational time versus error for both uniformly and gamma distributed random variables. Results in the bottom left corner of the graph are considered the best, as they represent low computational time and low error, whereas results in the top right hand corner are considered the worst, as they represent high computational time with high error.

When the random variables are uniformly distributed, the stochastic Galerkin method using third, fourth or fifth order expansions is clearly superior to Monte Carlo sampling, as they not only give a more accurate prediction but can also be calculated much more quickly. While using a first or second order stochastic Galerkin expansion is significantly faster than Monte Carlo sampling, the error is higher than simply using 500 Monte Carlo trials.

When the random variables are gamma distributed, the stochastic Galerkin method is also significantly faster than Monte Carlo sampling. However, it was not until a fourth order stochastic Galerkin expansion was used that the error was lower than that obtained with Monte Carlo sampling. It is interesting to note that increasing the order of the stochastic Galerkin expansion does not necessarily decrease the error. Increasing to a third order stochastic Galerkin expansion gave a higher error than a second order expansion in the case of gamma distributed random variables.

3.9 Discussion

In this chapter, uncertainty is introduced into the parameters and initial conditions of an SIR model using random variables with different probability distributions. As the parameters and/or initial conditions are now functions of random variables, the SIR model can no longer be solved using a single call to a numerical ODE solver. The mean and variance can be calculated using Monte Carlo sampling, but this can be computationally expensive. A superior method is the stochastic Galerkin method.

Even using low order expansions, it has been shown throughout this chapter that the stochastic Galerkin method produces accurate results when compared to those obtained through Monte Carlo sampling. Also, as the stochastic Galerkin method does not rely on random sampling, the stochastic Galerkin method is much quicker than Monte Carlo sampling.

Throughout this chapter, random variables with uniform, gamma and normal distributions are used to represent the uncertainty in the parameters and initial
Chapter 3. Applying the stochastic Galerkin method to the SIR model

Figure 3.18: Comparison of computational time against error of the stochastic Galerkin and Monte Carlo solutions for SIR model with uncertainty in $\beta$ and $\gamma$. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi_1$ and $\xi_2$ have either uniform distributions or gamma distributions. $I_{\text{initial}} = 0.01$. 

(a) Uniform Distribution

(b) Gamma Distribution
conditions. Random variables with other distributions will be discussed in later chapters. While the stochastic Galerkin method was shown to be accurate working with all these distributions, care must be taken when choosing the distributions of the random variables, as well as the values of the scaling parameters (such as $\beta_0$, $\gamma_0$, $a$ and $b$). Where possible, ensure $R_0 > 1$, for all values of the random variables. If this is not the case, predictions where no epidemic occurs will be averaged with predictions where an epidemic does occur. This means that the mean prediction is meaningless, as it does not accurately model either scenario.

Also, care must be taken when it is possible for $\beta$ or $\gamma$ to take negative values depending on the value of the random variables. It does not make real-world sense for these parameters to be negative. Including predictions where these parameters have negative values could affect the mean predictions obtained using the stochastic Galerkin method.

The work presented in this chapter provides several important extensions to the work already done by Roberts [12] and Chen-Charpentier and Stanescu [13] when applying the stochastic Galerkin method to an SIR model. Roberts only considered uncertainty in the basic reproduction number $R_0$. By considering the uncertainty in $\beta$ and $\gamma$ separately, much greater flexibility in the representation of the uncertainty is able to be achieved. The introduction of scaling parameters ($\beta_0$, $\gamma_0$, $a$ and $b$) is an improvement on the work done by Chen-Charpentier and Stanescu. By introducing scaling parameters, the deterministic system of ODEs resulting from the stochastic Galerkin method can be reused for different probability distributions for $\beta$ and $\gamma$. This is especially useful when the random variables are uniformly or normally distributed, as changing the scaling parameters can achieve any desired uniform or normal distribution for $\beta$ and $\gamma$. Uncertainty in the initial conditions of an SIR model was also investigated in this chapter whereas Roberts, as well as Chen-Charpentier and Stanescu, only considered deterministic initial conditions.

While using multiple uncertain parameters requires slightly more effort to set up, due to using multiple random variables, it allows for much greater analysis of the variance through the calculation of Sobol indices. The calculation of Sobol indices was not necessary in Roberts’ work, as only one uncertain parameter was considered, and while Chen-Charpentier and Stanescu did consider multiple parameters to have uncertainty, they did not calculate the Sobol indices for the SIR model.

The calculation of Sobol indices is important as it allows for greater understanding of the effect of the uncertainty each of the parameters has on the predictions from the epidemic model. As the Sobol indices are calculated directly from the stochastic Galerkin expansion, very few additional calculations are required. While not discussed in this chapter, the stochastic Galerkin method also allows for the uncertain parameters to be represented by random variables with different probability distributions. This is discussed further in Chapter 5.
Chapter 4

Issues with ‘blow-outs’

In the previous chapter, it was shown that the stochastic Galerkin method produced accurate results for the mean and variance when compared to results obtained from Monte Carlo sampling. However, this is not always the case. Depending upon the parameter values, and the type of probability distributions of the random variables, the solutions for \( S_i(t) \) and \( I_i(t) \) can ‘blow-out’ to large, unreasonable values. When this happens, the numerical integration scheme used to solve the deterministic system of ODEs from the stochastic Galerkin method (for example, the MATLAB function `ode45`) fails to find solutions for \( S_i(t) \) and \( I_i(t) \). When a blow-out occurs, the mean and variance cannot be calculated using the stochastic Galerkin method and a different method must be used.

The possibility of blow-outs when applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters is noted by Roberts [12] and Chen-Charpentier and Stanescu [13] when using normally distributed random variables. However, the blow-outs are not investigated further. A more detailed breakdown of the possible explanation for blow-outs given by Roberts, as well as Chen-Charpentier and Stanescu, is given at the end of Section 4.1.

As blow-outs are a significant drawback to using the stochastic Galerkin method, understanding why blow-outs occur is extremely important. In this chapter, ‘blow-outs’ in the stochastic Galerkin method will be first demonstrated and then investigated.

4.1 Example of a blow-out

Consider an SIR model with uncertainty only in the transmission rate, \( \beta \). The differential equations are given by

\[
\frac{dS}{dt} = -(\beta_0 + a\xi_1)SI
\]
\[
\frac{dI}{dt} = (\beta_0 + a\xi_1)SI - \gamma I,
\]
where uncertainty has been introduced into the transmission rate by letting 
\( \beta = \beta_0 + a\xi_1 \) where \( \beta_0 \) and \( a \) are positive constants and \( \xi_1 \) is a random variable with a known probability distribution. Figure 4.1 shows a comparison of the stochastic Galerkin solution with the solution obtained with Monte Carlo sampling using \( 10^5 \) trials with \( \beta_0 = 4, \gamma = 1, a = 0.24 \) and \( \xi_1 \) is a gamma distributed random variable with scale factor and shape parameter equal to one. From the figure, it can be seen that the two solutions are almost identical, which confirms that the stochastic Galerkin method is producing accurate results. However, if \( a \) is increased from 0.24 to 0.25, the solutions for \( S_i(t) \) and \( I_i(t) \) blow-out, as seen in Figure 4.2. For \( t < 1.5 \), the solutions for \( S_i(t) \) and \( I_i(t) \) appear reasonable. However, at approximately \( t = 2 \), the solutions for \( S_i(t) \) and \( I_i(t) \) quickly increase in magnitude and blow-out, heading off to \( \pm \infty \).

While a complete stochastic Galerkin solution could not be obtained due to the blow-out, Figure 4.3 shows a comparison of the part of the stochastic Galerkin solution that could be calculated, with the solution obtained with Monte Carlo sampling using \( 10^5 \) trials. From the figure, it can be seen that for \( t < 2 \), the mean solution obtained from the stochastic Galerkin solution is almost identical to that obtained using Monte Carlo sampling. For \( t > 2 \), the stochastic Galerkin solution could not be calculated due to the blow-out. Similarly, for \( t < 2 \), the variance calculated using the stochastic Galerkin method is extremely similar to that obtained using Monte Carlo sampling. However, at approximately \( t = 2 \) the blow-out occurs and the variance could not be calculated past this point.

One logical explanation for this blow-out is that due to introducing a random variable to account for the uncertainty in the transmission rate, the parameter can take unrealistic values or values that violate the real world conditions that it is trying to model. As \( \beta \) is the transmission rate, \( \beta \) must be positive. If \( \beta \) were negative for certain values of the random variable \( \xi_1 \), it would be a plausible explanation for the blow out. However, in this example, \( \xi_1 \) is gamma distributed. Therefore \( \beta \) will always be positive, regardless of the value of \( \xi_1 \).

Another possible explanation for this blow-out is that, due to introducing a random variable into the transmission rate, the value of \( R_0 \) could be less than one, for some values of \( \xi_1 \), but greater than one, for other values of \( \xi_1 \). As the dynamics of the epidemic model are vastly different depending on whether \( R_0 \) is less than or greater than one, this could cause a blow-out in the stochastic Galerkin method. However, since \( \xi_1 \) is gamma distributed, \( R_0 \) will always be greater than one, regardless of the value of \( \xi_1 \). Therefore, it is not the introduction of uncertainty that has caused a problem. As can be seen in Figure 4.3, the Monte Carlo solution could be calculated, so it is not a problem with introducing a random variable into the epidemic model.

It is also important to note that this blow-out is not caused by a numerical
Figure 4.1: Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. $P=5$, $\beta_0=4$, $\gamma=1$, $a=0.24$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{initial}=0.01$. 

(a) Mean

(b) Variance
Figure 4.2: Blow-out in the stochastic Galerkin method at approximately $t = 2$. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 0.25$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.01$. 
problem in the ODE solver. In an attempt to avoid the blow-out, several numerical ODE solvers available in MATLAB were used, but all returned the same results. Finally, a simple, self-written, forward Euler method was written in MATLAB and used to solve the deterministic ODEs, but it also resulted in a blow-out. The blow-out is caused by a problem with the stochastic Galerkin method and not the numerical solver used to solve the resulting deterministic ODEs from the stochastic Galerkin method.

There are only two known papers that refer to the problem of blow-outs when applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters. Roberts [12] notes that there could be problems when the random variable is normally distributed, as a normally distributed random variable can take any value on the real number line. When introducing uncertainty into the reproduction rate of a SIR model using the random variable $\theta$, Robert states:

\[
\text{If we took } \theta \text{ to be normally distributed, we would take } \phi_i(\theta) \text{ to be Hermite polynomials. However, this led to convergence problems, because Hermite polynomials are defined over the real line, and regions of the parameter space then correspond to different qualitative dynamics.}
\]

While this is a reasonable explanation when the random variable is normally distributed, blow-outs can occur even when the parameter values are always positive as shown earlier.

Chen-Charpentier and Stanescu [13] also raise possible problems when the uncertain parameters in their SIR and SIRS models ($a$, $b$ and $c$) are normally distributed. They state:

\[
\text{In the epidemic models considered the parameters } a, b \text{ and } c \text{ must be necessarily positive. Since the normal probability density function has support the whole real axis, it is not appropriate to consider these parameters as normally distributed. While for small variances simulations may still be performed, larger variances will necessarily lead to numerical instability. Compactly supported densities (i.e. uniform), or densities concentrated on the positive real axis (i.e. exponential), must be used for realistic problems.}
\]

However, once again, blow-outs have been shown to occur even when the parameter values are always positive. Neither paper offered any other explanation as to why blow-outs occur and no further analysis into blow-outs was mentioned in either paper.

It is extremely important that these blow-outs in the stochastic Galerkin method are investigated further, as they are a significant drawback to stochastic Galerkin method. Throughout this chapter, blow-outs in the stochastic Galerkin method are investigated. Firstly, to understand why blow-outs sometimes occur when using the
Figure 4.3: Blow-out in the stochastic Galerkin method at approximately $t = 2$. Comparison of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIR model with uncertainty in $\beta$. Blue are susceptible and red are infected. $P = 5$, $\beta_0 = 4$, $\gamma = 1$, $a = 0.25$, $\xi_1$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.01$. 
Chapter 4. Issues with ‘blow-outs’

4.2 Investigation of blow-outs

To better understand why the solutions for $S_i(t)$ and $I_i(t)$ blew-out (as seen in Figure 4.2), it would be helpful to analytically study the system of deterministic ODEs rather than just trying to solve them with a numerical solver. However, a fifth order expansion ($P = 5$), as used in Figures 4.1 and 4.2, uses twelve ODEs which is far too many for a detailed analytical analysis. Therefore, the order of the stochastic Galerkin expansion needs to be decreased so that there are fewer deterministic ODEs.

Reducing to a zero order expansion ($P = 0$) and keeping $\xi_1$ as a gamma distributed random variable with shape and scale parameters equal to one, the stochastic Galerkin method requires only two ODEs, which are given by

\[
\frac{dS_0}{dt} = - (\beta_0 + a) S_0 I_0 \\
\frac{dI_0}{dt} = (\beta_0 + a) S_0 I_0 - \gamma I_0.
\]

However, these are essentially just the original $SIR$ equations with $\beta$ replaced with its mean value as

\[
E[\beta] = E[\beta_0 + a \xi_1] = \beta_0 + a,
\]

where $E$ is the mean. Assuming appropriate values of the constants $\beta_0$, $a$ and $\gamma$ are chosen, this system of equations will never blow-out. Therefore, an analytic analysis of a zero order expansion ($P = 0$) is not helpful in investigating blow-outs in the stochastic Galerkin method. In order to investigate blow-outs, a minimum first order expansion ($P = 1$) is needed. However, this already requires four differential equations ($dS_0/dt$, $dI_0/dt$, $dS_1/dt$, $dI_1/dt$), which means an analytic analysis would still be difficult.

What is really needed is a first order stochastic Galerkin expansion that only produces two deterministic ODEs. With only two ODEs, an analytic analysis would be much easier to perform. Phase planes could easily be drawn and analysed. Although this is not possible with an $SIR$ model, it is possible using an $SI$ or $SIS$ model. Both of these models can be written using only one differential equation, which means a first order stochastic Galerkin expansion would only have two ODEs. As the $SIS$ model has more interesting dynamics than an $SI$ model, an $SIS$ model will be used to further investigate blow-outs in the stochastic Galerkin method. While examining an $SIS$ model’s blow-outs is unlikely to be helpful in predicting when the $SIR$ model will blow-out, it may help to explain why the blow-outs are occurring.
4.3 Blow-outs in the SIS model

To be more general, consider an SIS model with uncertainty in $\beta$ and $\gamma$, given by

$$\frac{dI}{dt} = (\beta_0 + a\xi)(1 - I) - (\gamma_0 + b\xi)I,$$  \hspace{1cm} (4.1)

where $\beta_0$, $\gamma_0$, $a$ and $b$ are positive constants and $\xi$ is a random variable with a known probability distribution.

In the previous section, blow-outs were shown to occur with uncertainty only in $\beta$. However, for a more thorough and complete analytical analysis, it will now be assumed that there is also uncertainty in $\gamma$. The simpler case of uncertainty only in $\beta$ can easily be considered using the same analytic analysis by simply letting $b = 0$. Also, note that in this SIS model, the random variable associated with $\beta$ and $\gamma$ is the same random variable whereas in the previous chapter, $\beta$ and $\gamma$ had independent random variables. While this is not a necessary condition to produce blow-outs in the SIS model, it does produce more interesting results. Gamma, uniform and normal distributions will be considered for the random variable $\xi$.

4.3.1 Gamma distribution

In this section, it will be assumed that the random variable, $\xi$, in Equation (4.1) is gamma distributed with shape and scale parameters equal to one. Therefore the probability density function for $\xi$ is given by $w(\xi) = e^{-\xi}$.

The first order stochastic Galerkin expansion for the fraction of infected individuals is given by

$$I(t, \xi) = \sum_{i=0}^{1} I_i(t) \Psi_i(\xi),$$  \hspace{1cm} (4.2)

where $\Psi_i(\xi)$ are the Laguerre polynomials, as their weight function is $e^{-\xi}$, which matches the probability density function of $\xi$. Substituting Equation (4.2) into Equation (4.1) and performing a Galerkin projection results in the following two deterministic ODEs for $I_0$ and $I_1$,

$$\frac{dI_0}{dt} = -(\beta_0 + a)(I_0)^2 + 2aI_0I_1 - (\beta_0 + 3a)(I_1)^2$$
$$+ (\beta_0 - \gamma_0 + a - b)I_0 - (a - b)I_1$$

$$\frac{dI_1}{dt} = a(I_0)^2 - (2\beta_0 + 6a)I_0I_1 + (2\beta_0 + 11a)(I_1)^2$$
$$- (a - b)I_0 + (\beta_0 - \gamma_0 + 3a - 3b)I_1.$$

(4.3)

Figure 4.4 shows a comparison of the stochastic Galerkin solution with the solution obtained from Monte Carlo sampling using $10^5$ trials when $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$ and $b = 1$. Despite using only a first order stochastic Galerkin expansion, the
Figure 4.4: Comparision of stochastic Galerkin (dashed lines) and Monte Carlo solutions (solid lines) for SIS model with uncertainty in $\beta$ and $\gamma$. Red are infected. $P = 1$, $\beta_0 = 4$, $\gamma_0 = 1$, $a = b = 1$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{\text{initial}} = 0.1$. 
Figure 4.5: Blow-out of the stochastic Galerkin solution for SIS model with uncertainty in $\beta$ and $\gamma$. Blue is $I_0(t)$ and red is $I_1(t)$. $P = 1$, $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$, $b = 1.3$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one. $I_{initial} = 0.1$.

mean solutions are almost identical. However, for $t > 2$, the variance obtained from the two different methods start to differ. This is not surprising as it is only a first order expansion. To obtain a more accurate solution for the variance, a higher order stochastic Galerkin expansion would need to be used. While the stochastic Galerkin solution for the variance is not accurate, it is still reasonable. If $b$ is now increased from 1 to 1.3, the stochastic Galerkin solution blows-out as seen in Figure 4.5. At $t \approx 2.8$, $I_0(t)$ quickly declines and drops below zero. However, as $I_0(t)$ is the mean fraction of infected individuals, $I_0(t) < 0$ does not make sense.

Clearly something has significantly changed in the dynamics of the ODEs (Equations (4.3)) for the stochastic Galerkin method to change from producing reasonably accurate results (for a first order expansion) when $b = 1$, to blow-outs when $b = 1.3$. To try to understand what has changed, phase planes can be produced with different values of $b$.

Figure 4.6 shows phase planes of Equations (4.3) when $b = 1$ and $b = 1.3$. When $b = 1$ (Figure 4.6(a)), there are four equilibrium points (A, B, C and E) and the stochastic Galerkin solution did not blow-out. When $b = 1.3$ (Figure 4.6(b)), there are now only two equilibrium points (C and E) and the stochastic Galerkin method blew-out. Therefore the number and type of equilibrium points present could be a determining factor in why there is a blow-out in the stochastic Galerkin solution when $b = 1.3$ but not when $b = 1$. 
Chapter 4. Issues with ‘blow-outs’

Figure 4.6: Phase planes for SIS model with uncertainty in $\beta$ and $\gamma$ with $P = 1$. Black lines are nullclines ($dI_0/dt = dI_1/dt = 0$). Dark green dashed lines are stable/unstable manifolds. Blue lines are sample solutions. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$, $\xi$ is a gamma distributed random variable with scale factor and shape parameter equal to one.
It is important to note that the arrows in the vector fields in Figure 4.6 have been normalised. This was done due to large differences in magnitudes throughout the vector field. If this was not done, some of the arrows, especially those in the bottom right-hand corner of the phase plane, would have been quite large while those near the equilibrium points would have been very short and hard to see. As it is the equilibrium points that are of interest, it is important to clearly see these arrows. Because of this normalisation, some of the arrows close to the stable and unstable manifolds may appear to cross the manifold. However, this is clearly not the case and simply a drawback of normalising the arrows.

Also, it is important to note that as there was no uncertainty in the initial conditions, the initial condition for $I_1$ must be zero. Therefore, all initial conditions occur on the horizontal axis such that $0 \leq I_0 \leq 1$. However, for clarity, Equations (4.3) were also solved backwards in time to make the phase-planes easier to analyse. If there was uncertainty in the initial conditions, $I_1$ could have an initial condition other than zero. However, this would also result in a different set of deterministic ODEs which would result in a different phase plane.

### 4.3.2 Analytically calculating equilibrium points

Rather than drawing phase planes, the equilibrium points of Equations (4.3) can be calculated analytically by letting $dI_0/dt = dI_1/dt = 0$ giving

$$A_1(I_0)^2 + B_1 I_0 I_1 + C_1(I_1)^2 + D_1 I_0 + E_1 I_1 = 0$$  \hspace{1cm} (4.4)

$$A_2(I_0)^2 + B_2 I_0 I_1 + C_2(I_1)^2 + D_2 I_0 + E_2 I_1 = 0$$  \hspace{1cm} (4.5)

where

$A_1 = -(\beta_0 + a)$ \hspace{1cm} $A_2 = a$

$B_1 = 2a$ \hspace{1cm} $B_2 = -(2\beta_0 + 6a)$

$C_1 = -(\beta_0 + 3a)$ \hspace{1cm} $C_2 = 2\beta_0 + 11a$

$D_1 = \beta_0 - \gamma_0 + a - b$ \hspace{1cm} $D_2 = -(a - b)$

$E_1 = -(a - b)$ \hspace{1cm} $E_2 = \beta_0 - \gamma_0 + 3a - 3b$.

Equation (4.5) can be rearranged to give

$$(I_1)^2 = -\frac{1}{C_2} \left[A_2(I_0)^2 + B_2 I_0 I_1 + D_2 I_0 + E_2 I_1 \right].$$  \hspace{1cm} (4.6)

Alternatively, Equation (4.5) can be rearranged to give

$$I_1 = \frac{1}{2C_2} \left[-(B_2 I_0 + E_2) \pm \sqrt{(B_2 I_0 + E_2)^2 - 4C_2(A_2(I_0)^2 + D_2 I_0)} \right].$$  \hspace{1cm} (4.7)
Substituting Equation (4.6) into Equation (4.4) and rearranging gives

\[
(A_1 C_2 - A_2 C_1)(I_0)^2 + (C_2 D_1 - C_1 D_2)I_0 + [(B_1 C_2 - B_2 C_1)I_0 + (C_2 E_1 - C_1 E_2)]I_1 = 0.
\]

(4.8)

Substituting Equation (4.7) into Equation (4.8) and rearranging gives

\[
I_0 \left[ (JL - F^2)(I_0)^3 + (KL + JM - 2FG)(I_0)^2 + ((E_2)^2L + KM + JN - 2FH - G^2)I_0 + ((E_2)^2M + KN - 2GH) \right] = 0,
\]

(4.9)

where

\[
F = B_2(B_1 C_2 - B_2 C_1) - 2C_2(A_1 C_2 - A_2 C_1)
\]
\[
G = E_2(B_1 C_2 - B_2 C_1) + B_2(C_2 E_1 - C_1 E_2) - 2C_2(C_2 D_1 - C_1 D_2)
\]
\[
H = E_2(C_2 E_1 - C_1 E_2)
\]
\[
J = (B_2)^2 - 4A_2 C_2
\]
\[
K = 2B_2 E_2 - 4C_2 D_2
\]
\[
L = (B_1 C_2 - B_2 C_1)^2
\]
\[
M = 2(B_1 C_2 - B_2 C_1)(C_2 E_1 - C_1 E_2)
\]
\[
N = (C_2 E_1 - C_1 E_2)^2.
\]

This gives a quartic equation in \( I_0 \). From Equation (4.9) it can be seen that \( I_0 = 0 \) is always a solution and therefore \( I_0 = 0 \) is always an equilibrium point of Equations (4.3) and corresponds to the disease free equilibrium. This then leaves the following cubic equation in \( I_0 \)

\[
(JL - F^2)(I_0)^3 + (KL + JM - 2FG)(I_0)^2 + ((E_2)^2L + KM + JN - 2FH - G^2)I_0 + ((E_2)^2M + KN - 2GH) = 0,
\]

(4.10)

which, for simplicity, can be rewritten as

\[
(I_0)^3 + a_1(I_0)^2 + a_2 I_0 + a_3 = 0,
\]

(4.11)

where

\[
a_1 = (KL + JM - 2FG)/(JL - F^2)
\]
\[
a_2 = ((E_2)^2L + KM + JN - 2FH - G^2)/(JL - F^2)
\]
\[
a_3 = ((E_2)^2M + KN - 2GH)/(JL - F^2).
\]

It is important to note that

\[
JL - F^2 = -32(2\beta_0 + 11a)(\beta_0^4 + 10a\beta_0^3 + 30a^2\beta_0^2 + 24a^3\beta_0 + 6a^4)
\]
and, as \( \beta_0 \) and \( a \) must be positive, \( JL - F^2 \neq 0 \). Therefore Equation (4.10) will always be a cubic equation and it will always be possible to evaluate the constants \( a_1 \), \( a_2 \) and \( a_3 \).

Using the formula for the solution of a cubic [62], the solutions for Equation (4.11), \( I_0^a \), \( I_0^b \), \( I_0^c \), are

\[
\begin{align*}
I_0^a &= W + T - \frac{a_1}{3} \\
I_0^b &= -\frac{W + T}{2} - \frac{a_1}{3} + i\sqrt{3}(W - T) \\
I_0^c &= -\frac{W + T}{2} - \frac{a_1}{3} - i\sqrt{3}(W - T)
\end{align*}
\]  
(4.12)

where

\[
\begin{align*}
Q &= \frac{3a_2 - (a_1)^2}{9} \\
R &= \frac{9a_1 a_2 - 27a_3 - 2(a_1)^3}{54} \\
W &= \sqrt[3]{R + \sqrt{Q^3 + R^2}} \\
T &= \sqrt[3]{R - \sqrt{Q^3 + R^2}}
\end{align*}
\]  
(4.13)

Therefore Equations (4.3) can have up to four real equilibrium points. The \( I_0 \) values for each of the equilibrium points are given by \( I_0 = 0 \) (the disease free equilibrium), \( I_0^a \), \( I_0^b \) and \( I_0^c \). The corresponding values for \( I_1 \) for each of the equilibrium points can be found by rearranging Equation (4.8) to give

\[
I_1 = -\frac{(A_1C_2 - A_2C_1)(I_0)^2 + (C_2D_1 - C_1D_2)I_0}{(B_1C_2 - B_2C_1)I_0 + (C_2E_1 - C_1E_2)}. 
\]  
(4.14)

Therefore the four equilibrium points for Figure 4.6(a) are

- **E**: \( I_0 = 0 \), \( I_1 = 0 \)
- **C**: \( I_0 = 0.0433 \), \( I_1 = -0.1252 \)
- **A**: \( I_0 = 0.6171 \), \( I_1 = 0.1039 \)
- **B**: \( I_0 = 0.5981 \), \( I_1 = 0.1756 \).

The stability of each of the equilibrium points can also be determined analytically. The Jacobian, \( J \), of Equations (4.3) is given by

\[
J = \begin{bmatrix}
2A_1I_0 + B_1I_1 + D_1 & B_1I_0 + 2C_1I_1 + E_1 \\
2A_2I_0 + B_2I_1 + D_2 & B_2I_0 + 2C_2I_1 + E_2
\end{bmatrix},
\]

which has the characteristic equation for its eigenvalues, \( \lambda \), given by

\[
\lambda^2 - b_1\lambda + b_2 = 0,
\]
where
\[
b_1 = 2A_1 I_0 + B_1 I_1 + D_1 + B_2 I_0 + 2C_2 I_1 + E_2
\]
\[
b_2 = (2A_1 I_0 + B_1 I_1 + D_1)(B_2 I_1 + 2C_2 I_1 + E_2)
- (B_1 I_0 + 2C_1 I_1 + E_1)(2A_2 I_0 + B_2 I_1 + D_2).
\]

Therefore the eigenvalues of \( J \), \( \lambda_1 \) and \( \lambda_2 \), are given by
\[
\lambda_1, \lambda_2 = \frac{1}{2} \left[ b_1 \pm \sqrt{(b_1)^2 - 4b_2} \right]. \tag{4.15}
\]

Using Equation (4.15), the eigenvalues for each of the equilibrium points in Figure 4.6(a) can be determined analytically, which also determines the stability of each of the equilibrium points. The four equilibrium points of Figure 4.6(a), along with their stability, are given below. When \( b = 1 \) (and the stochastic Galerkin method did not blow-out), there is one unstable equilibrium (E), two saddle points (B and C) and one stable equilibrium (A).

<table>
<thead>
<tr>
<th>Equilibrium Point</th>
<th>( I_0 )</th>
<th>( I_1 )</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>Unstable</td>
</tr>
<tr>
<td>C</td>
<td>0.0433</td>
<td>-0.1252</td>
<td>-3</td>
<td>2.9528</td>
<td>Saddle</td>
</tr>
<tr>
<td>A</td>
<td>0.6171</td>
<td>0.1039</td>
<td>-3</td>
<td>-1.6538</td>
<td>Stable</td>
</tr>
<tr>
<td>B</td>
<td>0.5981</td>
<td>0.1756</td>
<td>-3</td>
<td>1.6685</td>
<td>Saddle</td>
</tr>
</tbody>
</table>

The equilibrium points of Figure 4.6(b) can also be determined analytically and are given below. Therefore when \( b \) is increased from 1 to 1.3, there are now only two real equilibrium points. There is still the unstable equilibrium point at the origin (E) that corresponds to the disease free equilibrium. Also, one of the saddle points (C) is still present. However, the second saddle point and the stable equilibrium are no longer present.

<table>
<thead>
<tr>
<th>Equilibrium Point</th>
<th>( I_0 )</th>
<th>( I_1 )</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>2.8243</td>
<td>1.9757</td>
<td>Unstable</td>
</tr>
<tr>
<td>C</td>
<td>0.0301</td>
<td>-0.0820</td>
<td>-1.9775</td>
<td>2.7755</td>
<td>Saddle</td>
</tr>
</tbody>
</table>

Because the stable equilibrium is no longer present, the solutions for \( I_0(t) \) and \( I_1(t) \) quickly increase in magnitude, causing the stochastic Galerkin method to blow-out as seen in Figure 4.5. Therefore, the presence of the stable equilibrium point is crucial to the stochastic Galerkin method. Additionally, this confirms that it is not a problem with the numerical ODE solver and instead a problem with the stochastic Galerkin method.

If it could be determined when the stable equilibrium point is present, for different values of \( a \) and \( b \), it could be quickly determined when the stochastic Galerkin solution is able to be calculated and when it would blow-out.

The discriminant, \( D \), of Equation (4.10) (cubic equation in \( I_0 \)) is given by [62]
\[
D = Q^3 + R^2 \tag{4.16}
\]
where $Q$ and $R$ are defined in Equations (4.13). The discriminant can either be positive, negative or zero. If

- $D > 0$: one root is real and two are complex (conjugate pair)
- $D = 0$: all roots are real and at least two are equal
- $D < 0$: all roots are real and unique.

In order to have the stable equilibrium point present, and therefore not have blow-outs in the stochastic Galerkin method, the discriminant of Equation (4.10) must be negative. Figure 4.7 shows graphs of the discriminant of Equation (4.10) calculated analytically (using Equation (4.16)). Additionally, these graphs show where the stochastic Galerkin method blew-out numerically. From the graphs, it can be seen that the analytical range in which blow-outs were expected agrees very well with what was found numerically. It can also be seen $a = b = 1$, which produced reasonably accurate results earlier (Figure 4.4), is right on the edge of where the stochastic Galerkin solution is able to be calculated. By increasing $b$ to 1.3, the stochastic Galerkin method was pushed into a region in which it would blow-out.

While the analytical range in which the stochastic Galerkin method blew-out matched very well with numerical testing for $P = 1$, calculating the range of blow-outs for higher order expansions ($P > 1$) analytically is not feasible. Figure 4.8 shows where the stochastic Galerkin method numerically blew-out for the SIS model at higher order expansions. As the ranges of blow-outs vary greatly depending upon the order of the expansion, calculating the range of blow-outs for a low order expansion does not help in determining the range of blow-outs for higher order expansions. From the figures, it appears as though the range in which the stochastic Galerkin solution can be calculated becomes smaller as the order of the expansion increases. However, as it is not feasible to investigate higher order expansions analytically, it is not clear why the range in which the stochastic Galerkin solution can be calculated decreases in size as the order increases.

### 4.3.3 Uniform distribution

In this section, it will be assumed that the random variable in the SIS model (Equation (4.1)) is uniformly distributed on $[-1,1]$. Therefore the probability density function for $\xi$ is given by $w(\xi) = 1/2$.

The first order stochastic Galerkin expansion is given by

$$I(t, \xi) = \sum_{i=0}^{1} I_i(t) \Psi_i(\xi)$$

where $\Psi_i(\xi)$ are the Legendre polynomials, as their weight function is 1. Substituting Equation (4.17) into Equation (4.1) and performing a Galerkin projection results in
Figure 4.7: Discriminant ($D$) for SIS model with uncertainty in $\beta$ and $\gamma$. Black line is $D = 0$ (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method numerically blew-out. The stochastic Galerkin solution could be calculated at blue diamond when $a = b = 1$, whereas the red diamond is $a = 1$, $b = 1.3$, where the stochastic Galerkin method blew-out. Bottom graph has larger bounds on $a$ and $b$ compared to top graph. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi$ has a gamma distribution, with scale factor and shape parameter equal to one.
the following two ODEs for $I_0$ and $I_1$,

\[
\begin{align*}
\frac{dI_0}{dt} &= -\beta_0(I_0)^2 - \frac{2aI_0I_1}{3} - \frac{\beta_0(I_1)^2}{3} + (\beta_0 - \gamma_0)I_0 + \frac{(a-b)I_1}{3} \\
\frac{dI_1}{dt} &= -a(I_0)^2 - 2\beta_0I_0I_1 - \frac{3a(I_1)^2}{5} + (a-b)I_0 + (\beta_0 - \gamma_0)I_1.
\end{align*}
\] (4.18)

The equilibrium points of Equations (4.18) can be calculated by letting $dI_0/dt = 0$ and $dI_1/dt = 0$ giving

\[
\begin{align*}
A_1(I_0)^2 + B_1I_0I_1 + C_1(I_1)^2 + D_1I_0 + E_1I_1 &= 0 \\
A_2(I_0)^2 + B_2I_0I_1 + C_2(I_1)^2 + D_2I_0 + E_2I_1 &= 0
\end{align*}
\] (4.19) (4.20)
where, for simplicity,

\[
\begin{align*}
A_1 &= -\beta_0 & A_2 &= -a \\
B_1 &= -\frac{2a}{3} & B_2 &= -2\beta_0 \\
C_1 &= -\frac{\beta_0}{3} & C_2 &= -\frac{3a}{5} \\
D_1 &= \beta_0 - \gamma_0 & D_2 &= a - b \\
E_1 &= \frac{a - b}{3} & E_2 &= \beta_0 - \gamma_0.
\end{align*}
\]

By rewriting the equations in this form, the formulas for the equilibrium points, eigenvalues and discriminant derived in Section 4.3.2 remain the same and do not need to be rederived now that a uniform distribution is being considered. However, it is important to note that \(A_k\) and \(B_k\) where \(k = \{1, 2\}\) have been redefined from Section 4.3.2.

As the random variable is now uniformly distributed, it is possible for \(\beta\) and \(\gamma\) to be negative if \(a\) and \(b\) are not chosen carefully. If \(a > \beta_0\) then \(\beta\) could take negative values depending on the value of \(\xi\). Similarly, if \(b > \gamma_0\) then \(\gamma\) could take negative values.

The discriminant of Equation (4.10) can again be used to determine the number of real equilibrium points of Equations (4.18) and when the stable attractor will be present. If the stable attractor is present, the stochastic Galerkin solution will be able to be calculated and will not blow-out.

Figure 4.9 shows a graph of the discriminant of Equation (4.10), calculated analytically (using Equation (4.16)). Additionally, Figure 4.9 also shows where the stochastic Galerkin method numerically blew-out. In the figure, \(\beta_0 = 4\) and \(\gamma_0 = 1\). Therefore, when \(a > 4\), \(\beta\) could be negative depending on the value of \(\xi\). Similarly, when \(b > 1\), \(\gamma\) could be negative. Therefore, it would not be surprising if the stochastic Galerkin method blows-up if \(a > 4\) or \(b > 1\).

From Figure 4.9, it can be seen that the range of blow-outs calculated analytically using the discriminant agrees quite well with what was found numerically. However, there are two small regions where the range of blow-outs found numerically is different to what was calculated analytically.

When \(a = 1\) and \(b = 3.2\) (as well as the surrounding region), it was determined analytically, using the discriminant, that the stochastic Galerkin solution could be calculated. However, when trying to calculate the stochastic Galerkin solution numerically, it blew-out. To try to understand why the numerical result disagrees with the analytic result, a phase plane can be drawn. Figure 4.10 shows a phase plane of Equations (4.18) when \(a = 1\) and \(b = 3.2\). From the phase plane, it can be seen that there are four equilibrium points. The equilibrium points can be calculated analytically using Equations (4.12). The stability of these equilibrium points can also be determined analytically by finding the eigenvalues of the Jacobian (Equation (4.15)).
The four equilibrium points, as well as their stability, are given below.

<table>
<thead>
<tr>
<th>Equilibrium Point</th>
<th>$I_0$</th>
<th>$I_1$</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>1.7298</td>
<td>4.2702</td>
<td>Unstable</td>
</tr>
<tr>
<td>C</td>
<td>0.1823</td>
<td>0.3223</td>
<td>-1.7302</td>
<td>4.2110</td>
<td>Saddle</td>
</tr>
<tr>
<td>A</td>
<td>0.7850</td>
<td>-0.8451</td>
<td>-4.2322</td>
<td>-0.7495</td>
<td>Stable</td>
</tr>
<tr>
<td>B</td>
<td>0.7098</td>
<td>-0.9913</td>
<td>-4.2558</td>
<td>0.7498</td>
<td>Saddle</td>
</tr>
</tbody>
</table>

As the stable equilibrium point (A) is present, it would be expected that the stochastic Galerkin method would not blow-out. However, by reexamining the top and middle phase planes in Figure 4.10, it can be seen that not all sample solutions are able to reach the stable equilibrium point (A). When starting at $I_0 = 0.01$ and $I_1 = 0$ (the initial conditions used to produce Figure 4.9), the solution is not able to reach the stable equilibrium due to the saddle point at B. Therefore, if the initial fraction of infected is very small ($I_{\text{initial}} \approx 0$), despite the stable equilibrium point being present, the stochastic Galerkin method will blow-out.

Figure 4.11 shows the discriminant of Equation (4.10) and is similar to Figure 4.9; however, $I_{\text{initial}}$ has been increased from 0.01 to 0.20. For values near $a = 1$, $b = 3.2$, the discriminant now correctly predicts where the stochastic Galerkin method will numerically blow-out. Using a larger value of $I_{\text{initial}}$, the stochastic Galerkin solution is able to reach the stable equilibrium point and does not blow-out.
Chapter 4. Issues with ‘blow-outs’

Figure 4.10: Phase planes for SIS model with uncertainty in $\beta$ and $\gamma$ with $P = 1$. Middle and bottom phase plane shows enlarged sections of top phase plane. Black lines are nullclines ($dI_0/dt = dI_1/dt = 0$). Dark green lines are stable/unstable manifolds. Blue lines are sample solutions. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 1$, $b = 3.2$. $\xi$ is a uniformly distributed random variable on $[-1, 1]$. 

Figure 4.11: Discriminant $(D)$ for SIS model with uncertainty in $\beta$ and $\gamma$. Black line is $D = 0$ (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method numerically blew out. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi$ has a uniform distribution on $[-1, 1]$. $I_{\text{initial}} = 0.20$

There is still a region in Figure 4.11 where, analytically it was calculated that the stochastic Galerkin method would blow-out, but numerically it did not. This is the region around $a = 6$, $b = 0$. Figure 4.12 shows a phase plane of Equations (4.18) when $a = 6$ and $b = 0$. From the phase plane, it can be seen that there are now only two equilibrium points. The equilibrium points can be calculated analytically using Equations (4.12). The stability of these equilibrium points can also be determined analytically by finding the eigenvalues of the Jacobian (Equation (4.15)) and are given below.

<table>
<thead>
<tr>
<th>Equilibrium Point</th>
<th>$I_0$</th>
<th>$I_1$</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>-0.4641</td>
<td>6.4641</td>
<td>Saddle</td>
</tr>
<tr>
<td>A</td>
<td>0.4572</td>
<td>0.4572</td>
<td>-6.8339</td>
<td>-1.4240</td>
<td>Stable</td>
</tr>
</tbody>
</table>

The origin (E), which is the disease free equilibrium, is now a saddle point and the other equilibrium point (A) is a stable equilibrium. As the stable equilibrium is present, the stochastic Galerkin solution can be calculated and will not blow-out.

Up until now, it was assumed that if there were only two real equilibrium points (the discriminant is positive), the stable equilibrium would be not be present. It was assumed that the only equilibria that would be present are the disease free equilibrium and a saddle point. As the stable attractor was assumed to not be present, it was therefore assumed the stochastic Galerkin solution would blow-out.
However, in this case, there are only two real equilibrium points, but one of them is the stable equilibrium point. Therefore, the stochastic Galerkin solution is able to be calculated in the area near $a = 6, b = 0$ and will not blow-out despite only two real equilibrium points being present. This is because the stable equilibrium point is present despite the discriminant being positive.

Unfortunately, this means that the discriminant is not always a viable method for determining if the stochastic Galerkin method will blow-out. While in most cases, ensuring that the discriminant is negative would mean the stable attractor is present and therefore the stochastic Galerkin solution would not blow-out, this is not always the case.

Figure 4.13 shows the numerical range of blow-outs of the stochastic Galerkin method for higher order expansions. From the graphs, it can be seen that the range of blow-outs remains roughly the same for the different order expansions. This is very different to when the random variable was gamma distributed, where, as the order increased, the range of blow-outs changed significantly.

### 4.3.4 Normal distribution

In this section, it will be assumed that the random variable $\xi$ in the SIS model (Equation (4.1)) is normally distributed with zero mean and unit variance. There-
Chapter 4. Issues with ‘blow-outs’

Figure 4.13: Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIS model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi$ is a uniformly distributed random variable on $[-1, 1]$.

Therefore, the probability density function for $\xi$ is given by

$$w(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}.$$ 

The first order stochastic Galerkin expansion is given by

$$I(t, \xi) = \sum_{i=0}^{1} I_i(t) \Psi_i(\xi), \quad (4.21)$$

where $\Psi_i(\xi)$ are the Hermite polynomials, as their weight function is $\exp(-\xi^2/2)$.

Substituting Equation (4.21) into Equation (4.1) and performing a Galerkin projection results in the following two ODEs for $I_0$ and $I_1$,

$$\frac{dI_0}{dt} = -\beta_0(I_0)^2 - 2aI_0I_1 - \beta_0(I_1)^2 + (\beta_0 - \gamma_0)I_0 + (a - b)I_1 + (\beta_0 - \gamma_0)I_1.$$ 

(4.22)
The equilibrium points of Equations (4.22) can be calculated by letting $dI_0/dt = 0$ and $dI_1/dt = 0$ giving

$$A_1(I_0)^2 + B_1 I_0 I_1 + C_1(I_1)^2 + D_1 I_0 + E_1 I_1 = 0 \quad (4.23)$$

$$A_2(I_0)^2 + B_2 I_0 I_1 + C_2(I_1)^2 + D_2 I_0 + E_2 I_1 = 0 \quad (4.24)$$

where, for simplicity,

$$A_1 = -\beta_0 \quad A_2 = -a \quad \quad B_1 = -2a \quad B_2 = -2\beta_0$$

$$C_1 = -\beta_0 \quad C_2 = -3a \quad \quad D_1 = \beta_0 - \gamma_0 \quad D_2 = a - b$$

$$E_1 = a - b \quad E_2 = \beta_0 - \gamma_0.$$ 

Once again, by rewriting the equations in this form, the formulas for the equilibrium points, eigenvalues and discriminant derived in Section 4.3.2 remain the same and do not need to be rederived now that a normal distribution is being considered.

As the random variable is now normally distributed, even if $a$ and $b$ are very small, it is still possible for $\beta$ and $\gamma$ to take negative values. This is because a normally distributed random variable can take any value on the real number line. Therefore, blow-outs in the stochastic Galerkin method should be expected, especially when $a$ or $b$ is large. However, if $a$ and $b$ are small, the probability of a negative $\beta$ or $\gamma$ value is very small and it is hoped that the stochastic Galerkin solution will not blow-out.

Similarly to when the random variable was gamma and uniformly distributed, the discriminant of Equation (4.10) can be used to determine the number of equilibrium points of Equations (4.22). If the discriminant is negative, all four equilibrium points will be present, including the stable attractor. With the stable attractor present, the stochastic Galerkin solution should be able to be calculated.

Figure 4.14 shows a graph of the discriminant of Equation (4.10) calculated analytically using Equation (4.16). Additionally, Figure 4.14 also shows where the stochastic Galerkin method numerically blew-out. From the figure, it can be seen that the range of blow-outs calculated analytically agrees quite well with what was found numerically. However, this is one small region (around $a = 0.3$ and $b = 1.4$) where analytically it was predicted that the stochastic Galerkin solution could be calculated, but it blew-out numerically.

To determine why the analytic range of blow-outs slightly differs from the numerical range of blow-outs, phase planes can be drawn using parameter values in this region. Figure 4.15(a) shows a phase plane when $a = 0.3$ and $b = 1.2$. With these parameter values, it was analytically predicated that the stochastic Galerkin solution could be calculated and when it was tested numerically, a blow-out did not occur. From the phase plane, it can be seen that there are four equilibrium points...
Figure 4.14: Discriminant \((D)\) for SIS model with uncertainty in \(\beta\) and \(\gamma\). Black line is \(D = 0\) (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘\(\times\)’ shows where the stochastic Galerkin method numerically blow-out. \(\beta_0 = 4, \gamma_0 = 1\), \(\xi\) is a normally distributed random variable with zero mean and unit variance. \(I_{\text{initial}} = 0.01\)
which are given below.

<table>
<thead>
<tr>
<th>Equilibrium Point</th>
<th>$I_0$</th>
<th>$I_1$</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>2.1</td>
<td>3.9</td>
<td>Unstable</td>
</tr>
<tr>
<td>C</td>
<td>0.2313</td>
<td>0.2401</td>
<td>-2.1008</td>
<td>3.8245</td>
<td>Saddle</td>
</tr>
<tr>
<td>A</td>
<td>0.7642</td>
<td>-0.3038</td>
<td>-3.8368</td>
<td>-1.6617</td>
<td>Stable</td>
</tr>
<tr>
<td>B</td>
<td>0.5896</td>
<td>-0.5019</td>
<td>-3.8922</td>
<td>1.6631</td>
<td>Saddle</td>
</tr>
</tbody>
</table>

As the stable attractor is present, all solutions move towards it and the stochastic Galerkin solution can be calculated. However, if $b$ is increased from 1.2 to 1.5, it can be seen in Figure 4.15(b) that, while the stable attractor is still present, not all solutions can reach the stable attractor due to the saddle point (B). For very small values of $I_0$ ($I_0 \approx 0$), the stochastic Galerkin method will blow-out despite the presence of the stable attractor.

Figure 4.16 shows a graph of the discriminant of Equation (4.10), similar to Figure 4.14; however, $I_{\text{initial}}$ has been increased from 0.01 to 0.2. Now that $I_{\text{initial}}$ has been increased, the solution can now reach the stable equilibrium. This means that the stochastic Galerkin solution can be calculated and does not blow-out near $a = 0.3$, $b = 1.2$. The analytic range in which the stochastic Galerkin method can be calculated now agrees with what was found numerically.

Similar to when the random variable was uniformly distributed in the previous section, analytically calculating the discriminant is not enough to determine if the stochastic Galerkin solution can be calculated. Changing the initial fraction of infected individuals can cause the stochastic Galerkin solution to blow-out despite the presence of the stable attractor. However, in this case, it only changes the range in which the stochastic Galerkin solution blows-out very slightly.

Figure 4.17 shows the numerical range of blow-outs of the stochastic Galerkin method for higher order expansions. From the graphs, it can be seen that calculating the range of blow-outs for a low order expansion does not necessarily help in determining the range of blow-outs for higher order expansions. However, as a general rule, as the order increases, the range of blow-outs also increases.

### 4.4 Returning to SIR model

In the previous section, blow-outs in the SIS model were investigated. This was done as the SIS model could be written using a single differential equation, which allowed for an analytical analysis of the first order stochastic Galerkin solution. Using this analysis, it was found that the stochastic Galerkin solution would blow-out if the stable attractor was not present.

However, the original example of a blow-out in this chapter was for the SIR model. This section will investigate blow-outs in the SIR model. Consider a SIR
Figure 4.15: Phase planes for SIS model with uncertainty in $\beta$ and $\gamma$ with $P = 1$. Black lines are nullclines ($dI_0/dt = dI_1/dt = 0$). Dark green lines are stable/unstable manifolds. Blue lines are sample solutions. $\beta_0 = 4$, $\gamma_0 = 1$, $a = 0.3$. $\xi$ is a normally distributed random variable with zero mean and unit variance.
model with uncertainty in $\beta$ and $\gamma$. The differential equations for this model are given by

$$\frac{dS}{dt} = -(\beta_0 + a\xi_1)SI$$
$$\frac{dI}{dt} = (\beta_0 + a\xi_1) - (\gamma_0 + b\xi_2)I,$$

where $\beta_0$, $\gamma_0$, $a$ and $b$ are positive constants and $\xi_1$ and $\xi_2$ are independent random variables with known probability distributions. While it is too difficult to analytically calculate where blow-outs will occur, even for low order stochastic Galerkin expansions, the range of blow-outs can be determined numerically, similar to the $SIS$ model in the previous section. As before, gamma, uniform and normal distributions will be considered for the random variables.

Figure 4.18 shows where the stochastic Galerkin blew-out numerically when $\xi_1$ and $\xi_2$ are gamma distributed random variables with unit shape and scale parameters. As the random variables are gamma distributed, $\beta$ and $\gamma$ will always be positive. However, even for small values of $a$ and $b$, the value of $R_0$ could be less than one, or greater than one, depending on the values of $\xi_1$ and $\xi_2$.

Using a first order expansion, the stochastic Galerkin solution could be calculated.

Figure 4.16: Discriminant ($D$) for $SIS$ model with uncertainty in $\beta$ and $\gamma$. Black line is $D = 0$ (calculated analytically using Equation (4.16)). Blue ‘+’ is where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method numerically blew-out. $\beta_0 = 4$, $\gamma_0 = 1$, $\xi$ has a normal distribution with zero mean and unit variance. $I_{\text{initial}} = 0.2$.
Figure 4.17: Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIS model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated, while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$, $\xi$ is a normally distributed random variable with zero mean and unit variance.

when $a$ and $b$ were less than 0.8. However, when using higher order expansions, the range in which the stochastic Galerkin solution could be calculated quickly reduces. It is interesting to note that for second order expansions and higher, the stochastic Galerkin solution can blow out even if $a$ or $b$ is zero. This is what happened in the original example of a blow-out in the SIR model (Section 4.1). Even though there was only uncertainty in $\beta$ ($b = 0$), a blow-out still occurred.

Figure 4.19 shows where the stochastic Galerkin method blew-out numerically when $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. As the random variables are now uniformly distributed, it is possible for $\beta$ to be negative if $a > 4$ or $\gamma$ to be negative if $b > 1$.

When using either a first or second order expansion, as long as $a$ and $b$ are chosen such that $\beta$ and $\gamma$ remain positive, the stochastic Galerkin solution is able to be calculated. However, when using third and fourth order expansions, this is no longer the case. When $a = 3$ and $b = 0.5$, the stochastic Galerkin method still blows-out despite both $\beta$ and $\gamma$ remaining positive. It is interesting to note that the range of blow-outs for the third and fourth order expansions is very similar, with
the range of blow-outs only sightly larger for the fourth order expansion.

Figure 4.20 shows where the stochastic Galerkin method blew-out numerically when $\xi_1$ and $\xi_2$ are normally distributed random variables with zero mean and unit variance. When using a first order expansion, the stochastic Galerkin solution does not blow-out until $a$ and $b$ approach two. However, as the order of the expansion increases, the range in which the stochastic Galerkin solution can be calculated quickly narrows. As the random variables are normally distributed, they can take any real value. Therefore, even choosing small values of $a$ and $b$, it is still possible for $\beta$ or $\gamma$ to be negative.

### 4.5 Discussion

In this chapter, it has been shown that the stochastic Galerkin solution cannot always be calculated and can instead blow-out. While some researchers [12, 13] had noted the possibility of blow-outs when using normally distributed random
variables, no further analysis of blow-outs had been conducted. Additionally, it has been shown that even when the random variables are gamma distributed (and therefore the parameters are always positive), blow-outs can still occur.

Due to the large number of deterministic ordinary differential equations generated by the stochastic Galerkin method, it is usually infeasible to draw phase planes of the deterministic system of ODEs, especially at higher order expansions. This made it hard to determine why the stochastic Galerkin method sometimes blew-out. By using an SIS model to reduce the number of governing equations, and using a first order stochastic Galerkin expansion, it was found that the stochastic Galerkin method would blow-out if the stable attractor was not present. By calculating the discriminant, it was possible to analytically determine regions where the stochastic Galerkin method was likely to blow-out. However, there were regions where the stochastic Galerkin solution could be calculated but the discriminant predicted a blow-out. Additionally, by changing the initial fraction infected, the stochastic Galerkin could blow-out despite not changing the values of the scaling parameters $a$ and $b$. While there has been some success at predicting ranges of blow-outs in low

Figure 4.19: Ranges of blow-outs calculated numerically for the stochastic Galerkin method for the SIR model with uncertainty in $\beta$ and $\gamma$. Blue ‘+’ shows where the stochastic Galerkin solution could be calculated while red ‘×’ shows where the stochastic Galerkin method blew-out. $\beta = 4$, $\gamma = 1$. $\xi_1$ and $\xi_2$ are uniformly distributed random variables on $[-1, 1]$. $I_{initial} = 0.01$. 

<table>
<thead>
<tr>
<th>$P$</th>
<th>Ranges of Blow-Outs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1" alt="Range of Blow-Outs" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image2" alt="Range of Blow-Outs" /></td>
</tr>
<tr>
<td>3</td>
<td><img src="image3" alt="Range of Blow-Outs" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image4" alt="Range of Blow-Outs" /></td>
</tr>
</tbody>
</table>
order expansions, these ranges are usually unhelpful in determining blow-out ranges for higher order expansions. In general, the range in which the stochastic Galerkin solution could be calculated became smaller as the order of the expansion increased. However, it could not be determined why this is the case, or if this trend continues for higher order expansions.

Ideally, it needs to be determined if the stochastic Galerkin solution can be calculated before trying to implement the method. Generating the deterministic system of differential equations required for the stochastic Galerkin method can be time consuming, especially for high order expansions. Generating this system of ODEs is essentially wasted time and processing power, if another method eventually has to be used. More research needs to be done to help determine if the stochastic Galerkin solution can be calculated before starting the stochastic Galerkin method.
Chapter 5

Applying the stochastic Galerkin method to a data set from an epidemic

In Chapter 3, the stochastic Galerkin method was applied to the SIR epidemic model and was shown to produce accurate results when compared to results obtained through Monte Carlo sampling. The stochastic Galerkin method was also shown to give a significant speed increase over Monte Carlo sampling. However, in each application, the probability distributions of the uncertain parameters were simply ‘chosen’ rather than being derived. Additionally, no data set from an actual epidemic was considered.

In this chapter, a data set from an epidemic that went through a small boarding school will be investigated. Rather than simply assuming probability distributions for the uncertain parameters, individualised parameter distributions will be derived for the uncertain parameters from the data available. The stochastic Galerkin method will then be applied using the derived probability distributions and the predictions compared to the known data points.

The purpose of determining individual parameter distributions is to predict the possible effect of the same disease in a new population. This way, predictions of the time course of the epidemic can be made, along with ranges of uncertainty on a day-by-day basis.

There is limited literature on the stochastic Galerkin method being applied to data from an epidemic. Although, Roberts [12] applied the method to data from an influenza outbreak in New Zealand, only one parameter, $R_0$, was considered to be uncertain. Santonja and Chen-Charpentier [28] applied the method to obesity data, but due to limited data points, the distributions of the uncertain parameters were simply assumed to be uniform. The methodology presented in this chapter extends previous research by considering multiple parameters to have uncertainty, while deriving individualised probability distributions for these parameters. This
allows for much greater flexibility in the representation of the uncertainty, which is hoped will produce more accurate predictions.

5.1 Boarding house influenza epidemic

The data set that will be investigated in this chapter is from an influenza epidemic that went through a small boarding school in the North of England [73]. At the start of the new school term, students returned to the boarding school, where there were 763 male students enrolled. One of the students came back infected with influenza and over the following weeks, a small influenza epidemic spread through the boarding school.

When a student began showing symptoms, they were confined to bed. Because of this, accurate records of the number of infected students on each day of the epidemic were kept. In general, reporting of epidemics usually mention only the number of new infected cases each day/week rather than the actual number of infected individuals at that time. Only having the number of new infected cases each day makes the data much harder to work with and fit models to. However, using the boarding school data, the number of infected students on a given day is known. Also, as it is a boarding school, it can be assumed that the students enrolled were effectively isolated from the surrounding population. This makes it a unique and almost ideal data set to investigate.

While accurate records of the number of infected students each day were kept, the journal article [73] only contains a graph of the number of infected students on any given day. Because of this, the actual number of infected students had to be estimated from the graph. These estimates can be found in Table 5.1.

5.2 Finding the ‘best’ values for the parameters $\beta$ and $\gamma$

The first step in understanding the epidemic that went through the boarding school is to try to fit a compartment epidemic model to the known data points. To do this, a suitable compartment model must be chosen. As the students were able to recover from the disease, and the article did not refer to any of the students becoming reinfected after recovering from the disease, a recovered compartment is needed, which excludes simple $SI$ and $SIS$ models. Therefore the simplest model to attempt to fit to the data is the $SIR$ model. While an $SEIR$ model would also be a reasonable model to attempt to fit to the data, an $SEIR$ model would require fitting an additional parameter. Therefore, for simplicity, the $SIR$ model was chosen.

In order to fit the $SIR$ model to the known data, appropriate values for $\beta$ and $\gamma$ must be found, as well as initial conditions. From Table 5.1, only one student
Chapter 5. Applying the stochastic Galerkin method to a data set

was infected initially giving $I_{\text{initial}} = 1/763$. While it is possible that some of the students were immune to the disease, either through prior infection or vaccination, for simplicity, it is assumed that there are initially no students in the removed compartment and that all remaining students are susceptible to the disease giving $S_{\text{initial}} = 762/763$.

To determine appropriate values of $\beta$ and $\gamma$, a simple least squares error formula is used. The error associated with the SIR model’s predictions, using chosen values of $\beta$ and $\gamma$, compared to the known data points is given by

$$E_{\beta \gamma} = \sqrt{\sum_{k=0}^{14} \left[ I^\beta_{\gamma}(k) - I_D(k) \right]^2},$$

(5.1)

where $I^\beta_{\gamma}(k)$ is the fraction of infected students on day $k$ predicted by the SIR model with the values of $\beta$ and $\gamma$ and $I_D(k)$ is the actual fraction of infected students on day $k$ as given in Table 5.1.

The SIR model is then solved using different values of $\beta$ and $\gamma$ and the associated error, when compared to the known data points, is calculated using Equation (5.1). The MATLAB function fminsearch is used to find the values of $\beta$ and $\gamma$ that resulted in the smallest error. This occurred when $\beta \approx 1.665$ and $\gamma \approx 0.453$, giving $R_0 \approx 3.67$ and $E^{1.665}_{0.453} \approx 0.086$. The known data points, as well as this ‘best fit’ from the SIR model using the above values of $\beta$ and $\gamma$, can be seen in Figure 5.1. From the figure, it can be seen that the SIR model fits the known data points reasonably well, with the exception of the last three data points.

As this is a relatively well known data set, there are published values of $\beta$ and

<table>
<thead>
<tr>
<th>Day</th>
<th>Number Infected</th>
<th>Number Infected (Normalised)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.0013</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0.0039</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>0.0092</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>0.0328</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>0.0944</td>
</tr>
<tr>
<td>5</td>
<td>222</td>
<td>0.2910</td>
</tr>
<tr>
<td>6</td>
<td>282</td>
<td>0.3696</td>
</tr>
<tr>
<td>7</td>
<td>256</td>
<td>0.3355</td>
</tr>
<tr>
<td>8</td>
<td>233</td>
<td>0.3054</td>
</tr>
<tr>
<td>9</td>
<td>189</td>
<td>0.2477</td>
</tr>
<tr>
<td>10</td>
<td>123</td>
<td>0.1612</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>0.0917</td>
</tr>
<tr>
<td>12</td>
<td>25</td>
<td>0.0328</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>0.0144</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>0.0052</td>
</tr>
</tbody>
</table>

Table 5.1: Number of students infected with influenza at an English boarding school of 763 male students.
Figure 5.1: Influenza spreading within a small boarding school. Blue squares are known data points. Green line is the ‘best fit’ using an $SIR$ model with $\beta = 1.665$ and $\gamma = 0.453$.

$\gamma$. For example, Murray [74] and Poisot [75] find the parameter values that result in the best fit to be $\beta/763 \approx 2.18 \times 10^{-3}$ which gives $\beta \approx 1.66$. Murray finds $\gamma \approx 0.440$ whereas Poisot finds $\gamma \approx 0.441$. The small differences in parameter estimates for $\beta$ and $\gamma$ are most likely due to slight differences in approximating the data from the graph contained in the original article [73]. Different approximations of the data points would lead to slightly different ‘best’ values for the $SIR$ model.

For comparison purposes, an $SEIR$ model is also fitted to the boarding school data. The $SEIR$ model (as discussed in Section 2.1.5) has an additional compartment compared to the $SIR$ model for individuals that have been exposed to the disease, but are not yet infectious. The differential equations for the $SEIR$ model are given by

\[
\frac{dS}{dt} = -\beta SI
\]
\[
\frac{dE}{dt} = \beta SI - \alpha E
\]
\[
\frac{dI}{dt} = \alpha E - \gamma I
\]
\[
\frac{dR}{dt} = \gamma I.
\]

Similar to the $SIR$ model, it is assumed that there is initially one infected student
and the remainder of the students are susceptible to the disease. Equation (5.1) is again used, as well as the MATLAB function fminsearch, to find the parameter values that result in the minimum error. This occurs when $\beta \approx 1.804$, $\alpha \approx 15.487$ and $\gamma \approx 0.462$, which results in $R_0$ value of approximately 3.904 and an error of approximately 0.084. Therefore the SEIR has a slightly smaller error (0.084 compared to 0.086) when compared to the SIR model. The $\beta$ and $\gamma$ values are also slightly higher in the SEIR model (1.804 compared to 1.665 and 0.462 compared to 0.453 respectively) when compared to the SIR model.

Figure 5.2 shows the solution from the SEIR compared to known data points, as well as the solution from the SIR model for comparison. The two epidemic curves are very similar, with the only real difference in the peak of the epidemic. Both predict a peak at approximately the same time, but the the SEIR model predicts a slightly higher peak than the SIR model. It is interesting to note that similar to the SIR model, the SEIR model overestimates the tail of the epidemic, especially the last three data points.

It is also important to note the very large value of $\alpha$ ($\alpha = 15.487$) that was calculated as the best fit for the SEIR model. As $\alpha$ is the rate at which individuals leave the exposed compartment and move to the infected compartment, this large
value shows that the exposed compartment is not needed due to the very small amount of time individuals remain in the compartment. Therefore, the comparable predictions and errors of the SIR and SEIR models, as well as the large $\alpha$ value in the SEIR model, justify using the simplicity of the SIR model to model the boarding school epidemic rather than the SEIR model.

### 5.3 Finding plausible values for $\beta$ and $\gamma$

In the previous section, the ‘best’ values for $\beta$ and $\gamma$ were found by minimising the error in Equation (5.1). However, even though the error had been minimised, the SIR model (as well as the SEIR model) did not pass exactly through all the known data points, especially those towards the end of the epidemic.

Now let us assume that the influenza outbreak has escaped the relatively isolated confines of the boarding school and spread into the surrounding population. For example, the disease could have easily been spread by a teacher who works at the school, but does not live at the school. What information could be obtained from the boarding school epidemic data to help predict the likely course the epidemic will take in the surrounding population?

The most obvious place to start is using the ‘best’ values for $\beta$ and $\gamma$ from the boarding school epidemic. While these ‘best’ values would give a ‘best guess’ as to the likely course an epidemic might take outside the boarding house, this method gives no confidence intervals on other plausible outcomes. Additionally, the ‘best’ values modelled the tail of the epidemic rather poorly. If these parameter values were used to model the epidemic in the surrounding population, the same is likely to occur.

To better understand what could happen if the disease spread to the surrounding population, instead of simply using an SIR model with the values of $\beta$ and $\gamma$ that resulted in the smallest error, ranges of values for $\beta$ and $\gamma$ that have relatively low error also need to be considered. To begin, an error of less than 0.25 ($E_{\beta\gamma}^3 < 0.25$) was considered small and therefore the associated $\beta$ and $\gamma$ values were considered plausible values even though there may have been other parameter values that produce a smaller error. Note that the smallest error obtained was $E_{0.453}^{1.665} \approx 0.086$, so this new error threshold is approximately three times the minimum possible error. In later sections, other error thresholds are investigated.

In order to find pairs of plausible values for $\beta$ and $\gamma$, an approximate Bayesian computation method was used [75, 76]. To do this, suitable prior distributions for $\beta$ and $\gamma$ must be determined. As there was no prior knowledge of likely probability distributions for $\beta$ and $\gamma$, uniform distributions for both $\beta$ and $\gamma$ are used. Due to the real-world interpretations of $\beta$ and $\gamma$, these parameters must be positive. After some preliminary testing, it is found that no $\beta$ value greater than 3 returned an error
less than 0.25 and similarly no \( \gamma \) value greater than 1 returned an error less than 0.25. Therefore, the prior distribution for \( \beta \) is chosen to be a uniform distribution on \([0,3]\) and the prior distribution for \( \gamma \) is chosen to be a uniform distribution on \([0,1]\). Clearly neither parameter can be zero, but by choosing uniform distributions that started at zero, there is no possibility of overestimating the smallest possible values of the parameters.

Random values are then chosen for \( \beta \) and \( \gamma \) based upon their respective prior distributions. The \textit{SIR} model is then solved using these parameter values and the associated error calculated using Equation (5.1). If the error is less than 0.25, the pair of values is considered plausible and stored in a list. Otherwise, the parameter values are discarded and new values chosen for \( \beta \) and \( \gamma \) based upon their respective distributions. This process is repeated until \( 2 \times 10^5 \) plausible pairs of values are found. It is found that \( 2 \times 10^5 \) plausible pairs, when plotted on 2D plot, clearly outlined the probability distributions of \( \beta \) and \( \gamma \), while still being relatively quick to calculate. The plausible values for \( \beta \) and \( \gamma \) are then plotted as a histogram and can be seen in Figure 5.3. Finding more plausible pairs of values for \( \beta \) and \( \gamma \) would likely have given a smoother distribution but would have also increased the computation time.

From Figure 5.3, the general shape of the probability distributions for plausible values of \( \beta \) and \( \gamma \) can be seen. It is interesting to note that while values for \( \beta \) and \( \gamma \) are chosen from uniform distributions, the distributions for the plausible values of \( \beta \) and \( \gamma \) are not uniform. The minimum and maximum plausible values for \( \beta \) were approximately 1.44 and 1.94, respectively, which are well within the interval \([0,3]\) that was used as a prior. Similarly the minimum and maximum plausible values for \( \gamma \) were 0.33 and 0.66, which are also well within the interval \([0,1]\).

As the probability distributions did not resemble any of the standard probability distributions, they are approximated using polynomials. The histograms are first normalised (so that the area was equal to one) and replotted as a scatter plot. A range of different order polynomials, from third order up to seventh order, are then fitted to the data using the \textsc{Matlab} function \texttt{polyfit}. The results can be seen in Figure 5.4. From the figures it can be seen that with the exception of the third order polynomial, all of the polynomials are a good fit to the data. As an example, the fourth order polynomial approximating the \( \beta \) probability density distribution, \( w(\xi_1) \), is given by

\[
\begin{align*}
w(\xi_1) &= \begin{cases} 
-310.1(\xi_1)^4 + 2073.2(\xi_1)^3 - 5208.3(\xi_1)^2 \\
+5828.6\xi_1 - 2450.2 & \text{if } 1.4413 \leq \xi_1 \leq 1.9387 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

and the fourth order polynomial approximating the \( \gamma \) probability density function,
Chapter 5. Applying the stochastic Galerkin method to a data set

Figure 5.3: Histograms of plausible $\beta$ and $\gamma$ values ($E_{\gamma}^\beta < 0.25$) obtained for the boarding house epidemic using approximate Bayesian computation.
\( w(\xi_2) \), is given by

\[
\begin{aligned}
  w(\xi_2) &= \begin{cases} 
-2576.2(\xi_2)^4 + 5244.8(\xi_2)^3 - 4045.3(\xi_2)^2 \\
+1394.1\xi_2 - 176.5 & \text{if } 0.3339 \leq \xi_2 \leq 0.6553 \\
0 & \text{otherwise}
\end{cases}
\end{aligned}
\]

Determining analytic expressions for the probability distributions of plausible values for \( \beta \) and \( \gamma \) was the first step in applying to the stochastic Galerkin method in order to determine the mean prediction from the \( SIR \) model and its variance. The next step was to extend the stochastic Galerkin method to work with these probability distributions.

### 5.4 Deriving associated orthogonal polynomials

Before the stochastic Galerkin method could be applied, it was necessary to find the associated orthogonal polynomials for each of the probability distributions. As the probability distributions were non-standard and had to be approximated by polynomials, it is unlikely that the associated orthogonal polynomials would be known and would instead need to be derived. This was easily done using a Gram-Schmidt orthonormalisation method [77].

Consider deriving the orthogonal polynomials, \( \Psi_i(\xi) \), that have \( w(\xi) \) as their weight function on the interval \([a, b]\), where \( a \) and \( b \) are real numbers such that \( a < b \). While the weight function function, \( w(\xi) \), could be any non-negative function, in this case, it is also a probability density function for the random variable, \( \xi \), as the orthogonal polynomials are being derived to match the probability density functions derived in the previous section. This gives

\[
\int_a^b w(\xi)d\xi = 1.
\]

Therefore the zero order orthogonal polynomial is chosen to be \( \Psi_0(\xi) = 1 \) so that

\[
\langle \Psi_0(\xi), \Psi_0(\xi) \rangle = 1,
\]

where

\[
\langle \Psi_i(\xi), \Psi_j(\xi) \rangle = \int_a^b \Psi_i(\xi)\Psi_j(\xi)w(\xi)d\xi.
\]

The first order polynomial is then given by

\[
\Psi_1(\xi) = \frac{Y_1(\xi)}{\sqrt{\langle Y_1(\xi), Y_1(\xi) \rangle}},
\]
Figure 5.4: Approximating the probability distributions for $\beta$ and $\gamma$ with polynomials. Blue is a third order approximation, magenta is a fourth order approximation, black is a fifth order approximation, cyan is a sixth order approximation and red is a seventh order approximation.
where
\[ Y_i(\xi) = \xi - \langle \xi, \Psi_0(\xi) \rangle \Psi_0(\xi). \]

Using the Gram-Schmidt orthonormalisation method ensures that \( \Psi_1(\xi) \) satisfies
\[ \langle \Psi_0(\xi), \Psi_1(\xi) \rangle = 0 \text{ and } \langle \Psi_1(\xi), \Psi_1(\xi) \rangle = 1. \]

In general, the higher order orthogonal polynomials are derived using
\[ \Psi_i(\xi) = \frac{Y_i(\xi)}{\sqrt{\langle Y_i(\xi), Y_i(\xi) \rangle}} \]
where
\[ Y_i(\xi) = \xi^i - \sum_{k=0}^{i-1} \langle \xi^i, \Psi_k(\xi) \rangle \Psi_k(\xi). \]

This ensures that the orthogonal polynomials fit the criterion that
\[ \langle \Psi_i(\xi), \Psi_j(\xi) \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases} \]

## 5.5 Applying the stochastic Galerkin method

Now that probability density functions have been found for the plausible values of \( \beta \) and \( \gamma \), and the associated orthogonal polynomials have been derived using Gram-Schmidt orthonormalisation, the stochastic Galerkin method can be applied to determine the mean epidemic curve as well as its variance.

A third order stochastic Galerkin expansion \((P = 3)\) was applied to the SIR model with uncertainty in \( \beta \) and \( \gamma \) (see Equations (3.29)). The resulting deterministic ODEs were derived for each of the different polynomial orders (from third up to seventh) that were used to approximate the probability density functions of \( \beta \) and \( \gamma \). This was done to investigate what effect the different order polynomial approximations would have on the predictions obtained from the stochastic Galerkin method. The results can be seen in Figure 5.5. A third order stochastic Galerkin expansion was used as it has been shown in Chapter 3 to produce accurate results, while still being relatively quick to calculate.

From Figure 5.5, it can be seen that the different order polynomials have almost no effect on the final stochastic Galerkin predictions. It is important to note that there are five different stochastic Galerkin predictions present on Figure 5.5, each with its own color. However, as the predictions are so similar, the graph looks mostly red as this corresponds to the seventh order polynomial approximation and was plotted last. Even the third order polynomial, which did not approximate the \( \beta \) and \( \gamma \) probability distributions very well, still has an extremely similar stochastic Galerkin solution when compared to the other order polynomial solutions.

It is interesting to note that while the mean solution is similar to the ‘best fit’,
Figure 5.5: Stochastic Galerkin solutions to SIR model using plausible \( \beta \) and \( \gamma \) values where \( E_{\gamma}^{\beta} < 0.25 \). Blue squares are known data points. Green is the ‘best’ fit. Blue uses third order polynomial approximation, magenta uses fourth order, black uses fifth order, cyan uses sixth order and red uses seventh order. Solid lines are the mean solutions and dashed lines are \( \pm \) one standard deviation from the mean. (All stochastic Galerkin solutions are extremely similar with red plotted last).

the mean peak is lower than the ‘best fit’ peak. However, both peaks occur at approximately the same time. Also, while most of the known data points are within one standard deviation of the mean, the last three data points are outside this range.

By considering a range of plausible values for \( \beta \) and \( \gamma \), rather than simply the values that gave the smallest error, there is now a confidence interval on what could possibly happen if the disease were to start spreading in the population surrounding the boarding school.

### 5.6 Investigating other error thresholds

In the previous section, an error threshold of \( E_{\gamma}^{\beta} < 0.25 \) was used to determine a range of plausible values for \( \beta \) and \( \gamma \). In this section, two additional error thresholds are investigated: \( E_{\gamma}^{\beta} < 0.15 \) and \( E_{\gamma}^{\beta} < 0.35 \).

For each of the two new error thresholds, lists of plausible pairs of \( \beta \) and \( \gamma \) values must be obtained. As before, \( \beta \) values are randomly chosen from a uniform distribution on \([0, 3]\) and \( \gamma \) values are randomly chosen from a uniform distribution.
Chapter 5. Applying the stochastic Galerkin method to a data set

Figure 5.6: Histograms of plausible $\beta$ and $\gamma$ values obtained using approximate Bayesian computation for the three different error thresholds.

on $[0, 1]$. The SIR model is then solved using these parameter values and the associated error calculated using Equation 5.1. If the error is less than the error threshold, the $\beta$ and $\gamma$ values are considered plausible and added to the respective list, otherwise they are discarded. These plausible values were then plotted as a histogram and the results can be seen in Figure 5.6. Plausible values for $\beta$ and $\gamma$ using an error threshold of $E^{\gamma}_\beta < 0.25$ have also been included for comparison.

The distributions for $R_0$ are also plotted by calculating $\beta/\gamma$ for each plausible pair. The results can be seen in Figure 5.7. For comparison, the ‘best’ values of $\beta$ and $\gamma$ (calculated in Section 5.2) gave an $R_0$ value of approximately 3.67. As the error threshold increases, the distributions for $R_0$ become wider. It is interesting to note that the peak of the three distributions always occurs at an $R_0$ value less
than 3.67 (which was obtained using the ‘best’ values). While parameter values for $\beta$ and $\gamma$ were originally chosen from uniform distributions, the distributions for the $R_0$ are not uniform. They are also not symmetrical. The modal $R_0$ for each of the three error thresholds is on the left side of the distribution. This is especially the case for the largest of the three error thresholds. Again, this highlights the benefit of considering $\beta$ and $\gamma$ distributions as opposed to a single distribution for $R_0$.

The histograms are then normalised (such that their area is one) and replotted as scatter plots and the probability distributions are approximated using fifth order polynomials. While it was shown earlier that fourth order polynomials would have most likely sufficed, using a fifth order polynomial added relatively little computation time and could potentially approximate the new distributions more accurately than a fourth order polynomial. The probability distributions approximated with fifth order polynomials can be seen in Figure 5.8.

From the figure, it can been seen that the probability distributions for $\beta$ and $\gamma$ when $E_{\gamma}^\beta < 0.15$ are quite narrow. This is because the error threshold is only slightly larger than the minimum possible error so only $\beta$ and $\gamma$ values close to the ‘best’ values are within the error threshold. As the error threshold increases, first to $E_{\gamma}^\beta < 0.25$ and then to $E_{\gamma}^\beta < 0.35$, the probability distributions for $\beta$ and $\gamma$ become wider. As the probability distributions become wider, the heights of the probability distributions decrease. This is so that the area under each of the probability distributions remains equal to one. Therefore as the error threshold increases, the probability that the parameter values are equal, or very close, to the ‘best’ fit parameter values decreases.

Once the probability distributions for $\beta$ and $\gamma$ had been approximated by fifth order polynomials, the associated orthogonal polynomials are derived using the Gram-Schmidt orthonomalisation method. The stochastic Galerkin method is then applied to find the mean epidemic curve for each of the error thresholds as well as its variance. The results can be seen in Figure 5.9.

From the figure, it can be seen that the mean solution using the smallest error threshold, $E_{\gamma}^\beta < 0.15$, is very similar to the ‘best fit’ solution. This is because the small error threshold only allowed plausible values of $\beta$ and $\gamma$ that were close to the ‘best’ values. However, even when using this small error threshold, the mean peak is still lower than the ‘best fit’, although both peaks occur at approximately the same time. As the error threshold increases, the mean peak continues to decrease, while also occurring slightly earlier in the epidemic. However, this is only by a very small amount, with all peaks still occurring on day six of the epidemic.

It is interesting to note that the peaks of mean plus one standard deviation are very similar for all three error thresholds, but especially for $E_{\gamma}^\beta < 0.25$ and $E_{\gamma}^\beta < 0.35$. However, the peaks of mean minus one standard deviation are significantly different for the three error thresholds. As the different error thresholds
Figure 5.7: Distributions for $R_0$ for the three different error thresholds.
Figure 5.8: Approximating the probability distributions for $\beta$ and $\gamma$ using fifth order polynomials. Red is $E_\beta^\gamma < 0.15$, blue is $E_\beta^\gamma < 0.25$ and black is $E_\beta^\gamma < 0.35$. 
Figure 5.9: Stochastic Galerkin solutions ($P = 3$) to SIR model using plausible $\beta$ and $\gamma$ values. Red is $E_\beta < 0.15$, blue is $E_\beta < 0.25$ and black is $E_\beta < 0.35$. Blue squares are known data points. Green is the ‘best’ fit. Solid lines are the mean solutions and dashed lines are ± one standard deviation from the mean.

all have approximately the same upper confidence interval, it provides a solid prediction as to the worst possible scenario if the disease were to escape the boarding school and spread into the surrounding population. However, as the lower confidence interval changes significantly based upon what error threshold is used, it is much harder to predict the best case scenario assuming the disease started spreading in the surrounding population.

5.7 Determining plausible values for $\beta$ and $\gamma$

without approximate Bayesian computation

So far in this chapter, plausible values of $\beta$ and $\gamma$ have been found using approximate Bayesian computation. While this has worked well and provided interesting results, it is relatively time consuming to construct the list of plausible pairs of $\beta$ and $\gamma$ values. It would be much better to determine the probability distributions for $\beta$ and $\gamma$ without random sampling.

In order to better understand the error associated with different $\beta$ and $\gamma$ values,
Chapter 5. Applying the stochastic Galerkin method to a data set

Figure 5.10: Heat map of error associated with different $\beta$ and $\gamma$ values for an SIR model compared to known data points.

a heat map of the error is produced. This can be seen in Figure 5.10. The lowest error occurs at the ‘best’ values for $\beta$ and $\gamma$. However, moving away from these values causes the error increases. The plausible parameter values for each of the error thresholds form a simple closed shape. The smallest error threshold has the shape with the smallest area while the largest error threshold has the shape with the largest area.

The next step is to determine probability distributions for $\beta$ and $\gamma$ for each of the three error thresholds using only information about the closed shape. Figure 5.11 shows a plot of the plausible values for $\beta$ and $\gamma$ on a coarse grid for $E_{\beta\gamma} < 0.25$.

For simplicity, it is assumed that $\beta$ and $\gamma$ have independent probability distributions. While it can be seen from Figure 5.10 and Figure 5.11 that this is not strictly true, it is a reasonable assumption based upon the physical interpretations of $\beta$ and $\gamma$. The rate at which an infected individual infects other susceptibles ($\beta S$) is independent of an infectious person’s recovery rate ($1/\gamma$). While the rest of the this chapter assumes that the uncertain parameters are independent, Chapter 6 explores the possibility of the uncertain parameters no longer being independent.

From Figure 5.11, it can be seen that there are two plausible $\gamma$ values when $\beta = 1.45$. If $\beta$ is increased to 1.5, there are now four plausible $\gamma$ values. This process is continued for each plausible $\beta$ value and the results plotted as a histogram. This can be seen in Figure 5.12(a). A similar method is then used for plausible $\gamma$
values. From Figure 5.11, it can be seen that there are four plausible \( \beta \) values when \( \gamma = 0.35 \). If \( \gamma \) is increased to 0.4, there are now nine plausible \( \beta \) values. This process is continued and the results can be seen in Figure 5.12(b). Using this method, histograms of plausible \( \beta \) and \( \gamma \) values can be obtained for each of the three error thresholds without the need for approximate Bayesian computation. While sampling is still required, it is no longer ‘random’ sampling as only specific combinations of parameters need to be tested. Additionally, the number of parameter pairs that need to be tested is significantly decreased.

Figure 5.13 shows a plot of plausible \( \beta \) and \( \gamma \) values for each of the three error thresholds using a fine grid. Using these plots, histograms of plausible values for \( \beta \) and \( \gamma \) are obtained. These histograms are then normalised and replotted as a scatter plot. The distributions are then approximated using fifth order polynomials. The results can be seen in Figure 5.14.

As before, the probability distributions for \( \beta \) and \( \gamma \) when \( E_{\beta}^\gamma < 0.15 \) are quite narrow. As the error thresholds increase, the probability distributions become wider. By examining the closed shape of plausible values, the need to generate large lists of plausible values using approximate Bayesian computation has been eliminated. This allows plausible \( \beta \) and \( \gamma \) probability distributions to be calculated much faster.
Figure 5.12: Histograms of plausible $\beta$ and $\gamma$ values using $E_\gamma^\beta < 0.25$ on a course grid.
Figure 5.13: Plots of plausible values for $\beta$ and $\gamma$ using different error thresholds. Blue ‘+’ are less than the error threshold while red ‘x’ are higher than the error threshold.
Figure 5.14: Approximating the probability distributions for $\beta$ and $\gamma$ using fifth order polynomials. Red is $E_\beta^\gamma < 0.15$, blue is $E_\beta^\gamma < 0.25$ and black is $E_\beta^\gamma < 0.35$. 
Once the probability distributions for $\beta$ and $\gamma$ have been determined, the associated orthogonal polynomials are derived using the Gram-Schmidt orthonormalisation method. The stochastic Galerkin method is then applied and the results can be seen in Figure 5.15. Comparing this with Figure 5.9, whose results were obtained by approximate Bayesian computation, it can be seen that the two figures are extremely similar. This validates the approach of simply examining the closed shape to find the probability distributions of plausible $\beta$ and $\gamma$ values which allows the mean epidemic curves, as well as their variances, to be determined much faster than using approximate Bayesian computation.

### 5.8 Further speed increases

In the previous section, significant speed increases were obtained by not using approximate Bayesian computation and instead simply considering the closed shape that the uncertain parameters formed on a 2D plot. Each error threshold formed its own closed shape, with the smallest error threshold having the shape with the smallest area, and the largest error threshold having the shape with the largest area. However, there are many points that were tested to create the plots in Figure 5.13...
that did not need to be tested. For example, if it is known that a particular parameter pair lies within the closed shape, it does not need to be tested as it will have an error less than the threshold. Similarly, if a parameter pair is known to lie outside the closed shape, it also does not need to be tested as it will have an error greater than the error threshold. Therefore, only the border of the closed shape needs to be determined. All parameter pairs that fall within the border of the closed shape are considered plausible parameter values and all parameter pairs that fall outside the closed shape can be ignored as their error is larger than the threshold.

A routine was written in MATLAB that could find the border of the closed shape for any error threshold. The routine began by finding a parameter pair that had an error lower than the given error threshold and was therefore inside the closed shape. Using the same $\beta$ value, larger $\gamma$ values were chosen until the parameter pair gave an error larger than the threshold and was therefore outside the closed shape. Intermediate $\gamma$ values were then tested until the edge of the border was found. This method is similar to the bisection method used to find the root of an equation. The $\beta$ value was then increased (or decreased) and the process repeated. However, as it is a closed shape, the new $\gamma$ value that formed the border of the closed shape would be very close to the previous $\gamma$ value. Therefore, once a point on the border of the closed shape had been found, it is a relatively simple process to travel ‘along’ the border of the shape. The MATLAB routine for finding the border of the closed shape is given in Appendix B.

The border of the closed shape for the three error thresholds (0.15, 0.25 and 0.35) can be seen in Figure 5.16. From the figures, it is clear that far fewer parameter pairs needed to be tested, when only the border of the closed shape needs to be found. For example, Figure 5.16(b) required 3,411 parameter pairs to be tested. Using the same size grid, producing a figure similar to Figure 5.13(b) would require 701,701 parameter pairs to be tested. There are a few parameter pairs that were tested that fall well within, or well outside the border of the shape. These were used in the initial discovery of the border of the shape.

While only needing to find the border of the shape is a significant speed increase over testing all parameter pairs on a 2D plot, it is important to note that the probability distributions obtained for the uncertain parameters are the same. Assuming the same size grid is used, a plausible value of $\beta$ will have the exact same corresponding number of plausible $\gamma$ values regardless of the method used and therefore the histograms and scatter plots are the same. However, due to the significant speed increase of only needing to find the border, a finer grid size could be used. However, this is unlikely to change the distributions of the uncertain parameters significantly unless the original grid size was too coarse.
Figure 5.16: Border of plausible values for $\beta$ and $\gamma$ at three different error thresholds. Red x have an error higher than the threshold while blue + have an error less than the threshold.
Chapter 5. Applying the stochastic Galerkin method to a data set

5.9 Discussion

This chapter has looked at applying the stochastic Galerkin method to a data set from an epidemic. Rather than simply finding the parameter values that result in the lowest error when fitting an SIR model to the data, a range of parameter values were considered that had a low error, but not necessarily the lowest error.

Distributions for the uncertain parameters were first calculated using approximate Bayesian computation. Despite choosing uniform distributions for the prior distribution of the uncertain parameters, the resulting distributions were non-standard and were approximated using polynomials. Different order polynomials were trialled but had little to no effect on the final predictions. Therefore fifth order polynomials were used to approximate the probability distributions of the uncertain parameters throughout the rest of the chapter.

As the probability distributions of the uncertain parameters were non-standard, the stochastic Galerkin method was extended to work with uncertain parameters of any distribution. However, the associated orthogonal polynomials were unlikely to be known and, thus, were derived using a Gram-Schmidt orthonormalisation method. The stochastic Galerkin method was then applied and the mean prediction, as well as its variance, was determined for a range of error thresholds. As a range of plausible parameters values had been considered, confidence intervals were able to be determined for each of the error thresholds. This gave improved predictions about what could happen if the disease was to spread into the population surrounding the boarding school.

Several speed increases were also implemented. It was found that the plausible parameter values formed a simple closed shape on a 2D plot which eliminated the need for approximate Bayesian computation. Additionally, only the border of the closed shape needed to be found in order to determine the probability distributions of the uncertain parameters. This significantly reduced the number of parameter pairs that had to be tested, while not compromising the accuracy of the parameter distributions.

For this data set, it was found that an SEIR model provided little to no improvement over the predictions obtained from an SIR model. As the initial conditions were known, this left only two uncertain parameters in the SIR model. If an SEIR model were to be used, or the initial conditions were not known with uncertainty, there would be more than two uncertain parameters. This means that the speed increases that were obtained by simply finding the border of the closed shape may not be possible, as the uncertain parameter combinations can no longer be shown on a 2D plot. Approximate Bayesian computation may have to be used to determine probability distributions for the uncertain parameters, which would negate the speed increases obtained in this chapter.
Chapter 6

Dependent Distributions

So far through this thesis, it has been assumed that the random variables associated with the uncertain parameters within the epidemic models have had independent probability distributions. For example, in the previous chapter when looking at the epidemic that spread through the boarding school, it was assumed that while the values of $\beta$ and $\gamma$ contained uncertainty, the uncertainty in $\beta$ and $\gamma$ was uncorrelated and therefore the random variables associated with $\beta$ and $\gamma$ had independent probability distributions. This was a reasonable assumption based upon the real world properties that these parameters are attempting to model. The parameter $\beta$ models the ‘infectiousness’ of the disease and how fast it is spreading to new susceptibles, while $\gamma$ depends only on the time that an individual remains infectious. Therefore when probability distributions were calculated, it made sense to assume that the probability distributions were independent. However, this is not really the case.

Consider Figure 6.1 which shows an error heat map when fitting an $SIR$ model to the boarding house epidemic data (similar to the error plots from the previous chapter). When using an error threshold of 0.35, plausible $\beta$ values range from approximately 1.35 to 2.10, while plausible $\gamma$ values range from approximately 0.30 to 0.80. In the previous chapter, as the probability distributions for $\beta$ and $\gamma$ were considered independent, any plausible value for $\beta$ could be paired with any plausible value for $\gamma$. However, from the figure, it can be seen that if $\beta$ is chosen to be 1.4, the corresponding plausible values for $\gamma$ are now restricted and only range from approximately 0.33 to 0.52. As the selected value of $\beta$ narrows the plausible range of $\gamma$ values, it does not really make sense for $\beta$ and $\gamma$ to have independent probability distributions.

Within the limited literature available on applying the stochastic Galerkin method to model epidemics with uncertainty in the parameters, none of the current research considers that the random variables associated with the uncertain parameters could be dependent. Roberts [12, 26] and Hickson and Roberts [27] only consider one parameter within their models to contain uncertainty and therefore only require one random variable. This precludes the possibility of dependent random variables
This chapter looks at extending the stochastic Galerkin method to work with uncertain parameters that have dependent distributions. The epidemic that went through the boarding school is revisited, but the random variables associated with $\beta$ and $\gamma$ are no longer assumed to be independent.

### 6.1 Determining a dependent probability distribution for $\beta$ and $\gamma$

As the random variables associated with $\beta$ and $\gamma$ are no longer assumed to be independent, the general shape of the new, dependent probability distribution must be determined for the boarding school epidemic. An SIR model is once again used.
Similar to the previous chapter, it is assumed that the initial conditions are known with certainty and that there is initially one infected student and all other students are susceptible to the disease. Figure 6.2 shows an error plot, constructed using Equation (5.1), for the SIR model for different values of $\beta$ and $\gamma$. This plot shows the same data as Figure 6.1, but has been drawn as a 3D plot with error on the $z$-axis, rather than as a heat map. The lowest point on the surface (the smallest error) occurs when $\beta \approx 1.66$ and $\gamma \approx 0.45$, which are the ‘best fit’ values determined in the previous chapter. Moving away from these values, by either increasing or decreasing the values of $\beta$ or $\gamma$, increases the error.

When constructing a dependent probability distribution for $\beta$ and $\gamma$, it makes sense for the parameter values with the lowest error to also have the highest probability. In addition, parameter values that cause large errors should have low probabilities, as they are much less likely to accurately model the disease outbreak. Therefore, the general shape of the dependent probability distribution is simply determined by calculating $1/\hat{E}_{\beta,\gamma}^3$, where $\hat{E}_{\beta,\gamma}^3$ is defined in Equation (5.1), for different $\beta$ and $\gamma$ values. This is shown in Figure 6.3. The peak occurs at $\beta \approx 1.66$ and $\gamma \approx 0.45$.

While Figure 6.3 gives the general shape of the probability distribution, in order to apply the stochastic Galerkin method, an analytic form of the probability distribution must be determined. Based upon the general shape of the surface, it is
approximated by the function, \( Z(\beta, \gamma) \), which is given by

\[
Z(\beta, \gamma) = A_3 e^{-B_3(\beta-C_3)^2} e^{-D_3(\gamma-E_3)^2} + F_3
\]  

(6.1)

where \( A_3, B_3, C_3, D_3, E_3 \) and \( F_3 \) are constants.

To find appropriate values for the unknown constants, an error formula is required to determine how closely \( Z(\beta, \gamma) \) approximates the surface in Figure 6.3. Using 169 points of comparison (a 13x13 grid), the error between \( Z(\beta, \gamma) \) and the surface in Figure 6.3, \( Z_{\text{error}} \), is given by

\[
Z_{\text{error}} = \sum_{m=0}^{12} \sum_{n=0}^{12} \left[ Z(1.4 + 0.05m, 0.3 + 0.05n) - \left( E_\gamma^{1,4+0.05m} \right)^{-1} \right]^2
\]  

(6.2)

where \( E_\gamma^{1,4} \) is given in Equation (5.1). Note that Equation (5.1) was also used to generate Figure 6.1 and Figure 6.2.

Using the MATLAB function \texttt{fminsearch}, values of \( A_3, B_3, C_3, D_3, E_3 \) and \( F_3 \) are determined that minimise the error (Equation (6.2)). The values are

\[
\begin{align*}
A_3 &= 7.1082 \\
B_3 &= 29.6721 \\
C_3 &= 1.6815 \\
D_3 &= 75.9618 \\
E_3 &= 0.4677 \\
F_3 &= 2.5472
\end{align*}
\]
which gives an error of approximately 66.43. It is important to note that when using the MATLAB function \texttt{fminsearch} to find the minimum error, the parameter values found are highly dependent on the initial values used. Because of this, \texttt{fminsearch} was run multiple times with varying initial conditions. The parameter values listed above resulted in the lowest error found. However, it is possible that using different initial conditions, a different set of parameter values could be found that result in a slightly smaller error.

Substituting the parameter values listed above into $Z(\beta, \gamma)$ gives

$$Z(\beta, \gamma) = 7.1082e^{-29.6721(\beta-1.6815)^2}e^{-75.9618(\gamma-0.4677)^2} + 2.5472.$$  \hspace{1cm} (6.3)

A plot of $Z(\beta, \gamma)$ can be seen in Figure 6.4. It is interesting to note that $Z(\beta, \gamma)$ has its peak at $\beta = 1.6815$ and $\gamma = 0.4677$, which is slightly different to the peak in Figure 6.3 ($1/E$ plot) which has its peak at $\beta \approx 1.66$ and $\gamma \approx 0.45$.

In order to have the peak of the analytic equation occur at the same $\beta$ and $\gamma$ values as the peak in Figure 6.3, a new analytic equation of the form

$$Z_2(\beta, \gamma) = A_4e^{-B_4(\beta-1.6685)^2}e^{-D_4(\gamma-0.4529)^2} + F_4$$  \hspace{1cm} (6.4)

is chosen. Not only would this cause the peak of the analytic equation to occur at the same $\beta$ and $\gamma$ values as the peak in Figure 6.3, this equation also has two less parameters to fit compared to Equation (6.1).

Using \texttt{fminsearch}, values for $A_4$, $B_4$, $D_4$ and $F_4$ are found that minimise the error between Equation (6.4) and the surface in Figure 6.3. The values are

$$A_4 = 7.1968 \hspace{1cm} B_4 = 31.7632$$
$$D_4 = 84.8105 \hspace{1cm} F_4 = 2.6284.$$  

However, the error of this simplified analytic form (calculated using Equation (6.2)) is approximately 79.35 compared to an error of 66.43 when using $Z(\beta, \gamma)$ (Equation (6.3)). Therefore forcing the peak to occur at the ‘best’ parameter values actually increased the error.

In another attempt to more closely match the analytic equation to the surface in Figure 6.3, an analytic equation of the form

$$Z_3(\beta, \gamma) = A_5e^{B_5 \beta^2 + C_5 \beta + D_5 e^{E_5 \gamma^2} + F_5 \gamma + G_5} + H_5$$  \hspace{1cm} (6.5)

was chosen. While $Z_3(\beta, \gamma)$ is essentially the same as $Z(\beta, \gamma)$, it is hoped that \texttt{fminsearch} may find parameter values that result in a lower error than those found using $Z(\beta, \gamma)$. The parameter values that result in the minimum error, as found by
Figure 6.4: Analytic approximation to $1/E_7^β$ plot (Figure 6.3) using Equation (6.3).

The values of $A_5$, $B_5$, $C_5$, $D_5$, $E_5$, $F_5$, $G_5$, and $H_5$ are:

\[
\begin{align*}
A_5 &= 11.1642 & B_5 &= -29.6699 \\
C_5 &= 99.7806 & D_5 &= -95.7806 \\
E_5 &= -75.9610 & F_5 &= 71.0563 \\
G_5 &= -5.4054 & H_5 &= 2.5471 \\
\end{align*}
\]

and these give an error of approximately 66.43, which matches the error when using $Z(\beta, \gamma)$. While $Z_3(\beta, \gamma)$ did not result in a smaller error, it did indicate that this is most likely the smallest error possible using an exponential fit. As the errors for $Z(\beta, \gamma)$ and $Z_3(\beta, \gamma)$ are the same, it was decided to simply use $Z(\beta, \gamma)$ to analytically represent the surface in Figure 6.3 as it is simpler to use.

Now that an analytic approximation for the general shape of the probability distribution has been determined, the next step is to choose limits for $\beta$ and $\gamma$ and to normalise $Z(\beta, \gamma)$, so that the volume underneath the surface is equal to one (which is a requirement of a probability density function). As $F_3 > 0$, limits need to be placed on the ranges of $\beta$ and $\gamma$ values so that the volume under the surface of $Z(\beta, \gamma)$ remains finite. The values of $\beta$ were restricted such that $1.4 \leq \beta \leq 2.0$ and the values of $\gamma$ were restricted such that $0.3 \leq \gamma \leq 0.9$. These ranges ensure that the height of the surface at the boundaries is very close to $F_3$ and therefore $\beta$ and $\gamma$ values outside these ranges have very low probabilities, as seen in Figure 6.4. These
ranges are also similar to the plausible ranges for $\beta$ and $\gamma$ calculated in Chapter 5 ($1.35 \leq \beta \leq 2.1$ and $0.3 \leq \gamma \leq 0.8$) using the largest error threshold, $E_\beta^\gamma < 0.35$. Using these ranges, the volume, $V$, under the surface of $Z(\beta, \gamma)$ can be calculated.

$$V = \int_{0.3}^{0.9} \int_{1.4}^{2.0} Z(\beta, \gamma) d\beta d\gamma \approx 1.3680.$$  

Dividing $Z(\beta, \gamma)$ by 1.368 normalises the dependent probability density function, $w(\xi_1, \xi_2)$, for $\beta$ and $\gamma$ and is given by

$$w(\xi_1, \xi_2) = 5.1960 e^{-29.6721(\xi_1 - 1.6815)^2} e^{-75.9618(\xi_2 - 0.4677)^2} + 1.8620,$$

where $\xi_1$ is the random variable that represents the uncertainty in $\beta$ and $\xi_2$ is the random variable that represents the uncertainty in $\gamma$.

### 6.2 Applying the stochastic Galerkin method

Now that an analytic form of the dependent probability distribution for $\beta$ and $\gamma$ has been determined, the stochastic Galerkin method can be applied to determine the mean prediction and its variance. To do this, $S$ and $I$ (the fraction of susceptible
and infected individuals respectively) are rewritten in the form

\[
S(t, \xi_1, \xi_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} S_{ij}(t) \Psi_{ij}(\xi_1, \xi_2)
\]

\[
I(t, \xi_1, \xi_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} I_{ij}(t) \Psi_{ij}(\xi_1, \xi_2),
\]

where \(\Psi_{ij}(\xi_1, \xi_2)\) are the two dimensional orthogonal polynomials whose weight function is \(w(\xi_1, \xi_2)\). As \(w(\xi_1, \xi_2)\) is a function of both the random variables, \(\xi_1\) and \(\xi_2\), and cannot be separated into the product of functions of only one of the random variables, the associated two dimensional orthogonal polynomials must be used. It is no longer possible to simply use the product of one dimensional orthogonal polynomials, as done in the previous chapter.

As \(w(\xi_1, \xi_2)\) is not one of the standard probability distributions, the associated orthogonal polynomials are unlikely to be known and must be derived. This is done using the Gram-Schmidt orthonormalisation method (similar to the previous chapter), which gives the following orthogonal polynomials:

\[
\Psi_{00}(\xi_1, \xi_2) = 1
\]

\[
\Psi_{10}(\xi_1, \xi_2) = 6.3274\xi_1 - 10.7231
\]

\[
\Psi_{01}(\xi_1, \xi_2) = 6.2377\xi_2 - 3.4786
\]

\[
\Psi_{20}(\xi_1, \xi_2) = 39.4217(\xi_1)^2 - 133.9320\xi_1 + 112.7706
\]

\[
\Psi_{11}(\xi_1, \xi_2) = 37.1784\xi_1\xi_2 - 22.5293\xi_1 - 63.4688\xi_2 + 38.4214
\]

\[
\Psi_{02}(\xi_1, \xi_2) = 39.9600(\xi_2)^2 - 47.5560\xi_2 + 13.0661.
\]

While it is a relatively simple process to generate the two dimensional orthogonal polynomials, with weight function \(w(\xi_1, \xi_2)\), using the Gram-Schmidt orthonormalisation method for low orders, it can be computationally expensive, and possibly ill-conditioned, for higher orders.

Substituting Equations (6.6) into the SIR model and performing a Galerkin projection (using the above orthogonal polynomials) gives a deterministic system of differential equations. It was hoped that by solving this system of differential equations, the mean and variance of the boarding house epidemic could have been calculated using the new, dependent distribution for \(\beta\) and \(\gamma\). However, despite the amount of work put in so far, the solutions ‘blew-out’, similar to what was shown in Chapter 4.

In an attempt to calculate usable predictions from the stochastic Galerkin method, the ranges for \(\beta\) and \(\gamma\) were slightly adjusted. Firstly, the range of \(\beta\) was modified so that \(1.45 \leq \beta \leq 1.95\) (rather than \(1.4 \leq \beta \leq 2.0\)) and the range of \(\gamma\) was left as \(0.3 \leq \gamma \leq 0.9\). While this appears to be a simple change, the majority of steps in
the stochastic Galerkin method had to be redone.

Firstly, the volume under the analytic curve was recalculated so that the new probability distribution could be normalised. The new volume was given by

\[ V = \int_{0.3}^{0.9} \int_{1.45}^{1.95} Z(\beta, \gamma) \, d\beta \, d\gamma \approx 1.1993. \]

Dividing Equation (6.3) by 1.1993 gave the new dependent probability density function, \( w_2(\xi_1, \xi_2) \), for \( \beta \) and \( \gamma \)

\[ w_2(\xi_1, \xi_2) = 5.9269e^{-29.6721(\xi_1 - 1.6815)^2}e^{-75.9618(\xi_2 - 0.4677)^2} + 2.1238. \quad (6.7) \]

As the probability density function changed, the associated orthogonal polynomials had to be rederived. The first few two dimensional orthogonal polynomials with \( w_2(\xi_1, \xi_2) \) as their weight function are given by

\[
\begin{align*}
\Psi_{00}(\xi_1, \xi_2) & = 1 \\
\Psi_{10}(\xi_1, \xi_2) & = 7.4839\xi_1 - 12.6853 \\
\Psi_{01}(\xi_1, \xi_2) & = 6.3156\xi_2 - 3.4951 \\
\Psi_{20}(\xi_1, \xi_2) & = 56.4927(\xi_1)^2 - 191.8947\xi_1 + 161.9481 \\
\Psi_{11}(\xi_1, \xi_2) & = 44.2482\xi_1\xi_2 - 26.8126\xi_1 - 75.5100\xi_2 + 45.7110 \\
\Psi_{02}(\xi_1, \xi_2) & = 40.3072(\xi_2)^2 - 47.8992\xi_2 + 13.1527.
\end{align*}
\]

Once the orthogonal polynomials had been re-derived, the deterministic system of ordinary differential equations also had to be re-derived. While the form of the equations was similar to the system of ODEs derived earlier, the coefficients were determined by inner products, involving the orthogonal polynomials and probability density function, which were now different. By narrowing the range of \( \beta \) values, the stochastic Galerkin solution could now be calculated and no longer ‘blew out’. The solution can be seen in Figure 6.6.

As narrowing the range of possible \( \beta \) values solved the blow-out problem, an attempt to find a stochastic Galerkin solution by slightly altering the range of possible \( \gamma \) values was made. The range of possible \( \gamma \) values was restricted such that \( 0.3 \leq \gamma \leq 0.8 \) (rather than \( 0.3 \leq \gamma \leq 0.9 \)) while using the original range of \( \beta \) values \( (1.4 \leq \beta \leq 2.0) \).

As before, the volume under the analytic curve had to be recalculated so that the probability density function could be normalised. The new volume was given by

\[ V = \int_{1.4}^{2.0} \int_{0.3}^{0.8} Z(\beta, \gamma) \, d\beta \, d\gamma \approx 1.2152. \]

Dividing Equation (6.3) by 1.2152 gave the new dependent probability density func-
Figure 6.6: Stochastic Galerkin solutions to SIR model when $\beta$ and $\gamma$ have a dependent probability distribution. Blue squares are known data points. Green is the ‘best’ fit. Solid lines are the mean solution and dashed lines are $\pm$ one standard deviation from the mean. Red used a narrowed $\beta$ range, while blue used a narrowed $\gamma$ range.

It is interesting to note that Equation (6.8) is very similar to Equation (6.7) as the two volumes were similar ($1.2152$ compared to $1.1993$). However, despite the similar probability density functions, the associated orthogonal polynomials still had to be recalculated. The first few are given by

$$w_3(\xi_1, \xi_2) = 5.8494e^{-29.6721(\xi_1-1.6815)^2}e^{-75.9618(\xi_2-0.4677)^2} + 2.0961.$$

By narrowing the range of $\gamma$ values slightly, the stochastic Galerkin solution could
now be calculated and no longer blew-out. The predictions can be seen in Figure 6.6, as well as the predictions from the stochastic Galerkin method found when the range of \( \beta \) values was narrowed.

From the graph, it can be seen that the peak of both mean stochastic Galerkin predictions occur at approximately the same time as the ‘best fit’ prediction. However, both stochastic Galerkin predictions significantly underestimate the peak of the epidemic, with the prediction obtained using a narrowed \( \gamma \) range having a slightly higher peak than the the prediction obtained using a narrowed \( \beta \) range. According to the known data points, the peak of the epidemic has approximately 37% of the students infected whereas the mean stochastic Galerkin predictions only predict a peak of 26%, using the narrowed \( \beta \) range, and 29%, using the narrowed \( \gamma \) range. By day ten, the two mean predictions from the stochastic Galerkin method, as well as the ‘best fit’ prediction, almost match the known data point. From this point on, all three predictions are very similar, but all overestimate the tail of the epidemic.

While most of the data points do lie within one standard deviation of the mean predictions, the standard deviations are quite large, especially around the peak of the epidemic. Using the standard deviations as a confidence interval, using the narrowed \( \beta \) range, the peak of the epidemic could have been anywhere between 14% and 39% of the students infected. Similarly, determining a confidence interval using the narrowed \( \gamma \) range, anywhere between 20% and 41% of the students could be infected at the peak of the epidemic. This is a wide range of scenarios, which is disappointing, as it was hoped that by assuming the uncertain parameters were dependent, the predictions would have more closely matched the data points.

### 6.3 Discussion

This chapter has shown that the stochastic Galerkin method can still be used when the random variables associated with the uncertain parameters are no longer independent. However, one of the main difficulties is trying to find an analytic approximation for the probability distribution. In this chapter, it was assumed that the general form of the probability distribution was proportional to \( 1/E^\beta \gamma \) (Equation (5.1)). However, other forms for the general shape of the probability distribution could be considered.

By looking at the general shape of the probability distribution, a likely analytic equation to model the surface was chosen. Several variations were considered and the one with the lowest error was eventually used. However, there could be other analytic forms that also accurately model the surface. For example, Lagrange interpolating polynomials could be used to approximate the surface. This was attempted, but suffered from large fluctuations of the polynomials between interpolating points, even when using low order interpolating polynomials. However, this method may
be suitable for other probability distributions that arise from different data sets.

There are also several different ways the analytic function could have been nor-
malised such that the volume underneath the surface was equal to one. In this
chapter, the initial volume was found using numerical integration and the func-
tion divided by the volume. In each case, the volume was already close to one
(1.3680, 1.1993 and 1.2152), but this may not be the case for other data sets. The
analytic function could also have been normalised by lowering the value of $F_3$ (Equa-
tion (6.1)). This would have the additional advantage of further reducing the prob-
ability of $\beta$ and $\gamma$ values that were far away from the peak. Alternatively, the ranges
of the uncertain parameters could be restricted in order to normalise the volume.
Whichever method is used to normalise the volume, care must be taken to ensure
that the resulting probability density function is always positive.

Throughout this chapter, blow-outs in the stochastic Galerkin method have been
shown to be a problem. The work done to calculate the orthogonal polynomials and
derive the deterministic system of equations for the stochastic Galerkin method is
all wasted when it is ultimately found that a blow-out has occurred. However, it was
possible to reuse the analytic equation that modelled the probability distribution.
Simply narrowing the range of possible $\beta$ or $\gamma$ values resolved the problem of blow-
outs, but this still required both the orthogonal polynomials and the deterministic
system of differential equations to be recalculated. Even when only slightly modify-
ing the ranges of the uncertain parameters, this caused differences in the predictions
obtained. If the ranges of the uncertain parameters had to be changed significantly
in order to obtain ‘sensible’ results from the stochastic Galerkin method, this could
potentially have a large impact on the final predictions.
Chapter 7

Using the stochastic Galerkin method as a predictive tool

In the previous two chapters, the stochastic Galerkin method was used to make predictions about what could happen if the influenza virus escaped the boarding school and entered the population in the surrounding area. This was done using all available data points from the boarding school outbreak. However, these predictions would not be of use to those at the boarding school.

Real-time forecasting of epidemics is of great importance. There is extensive literature, using a range of methods, for predicting the likely course an epidemic will take with limited initial data [78, 39]. As new data becomes available, predictions are able to be updated and refined. Some example prediction methods include fitting parameters to compartment models [79], Bayesian particle filters [80, 81], Monte Carlo maximum likelihood analysis [82], discrete time stochastic models [83] as well as individual based models [84, 85].

As data was collected on each day of the boarding school epidemic, could predictions have been made during the epidemic to determine the likely course the epidemic would take? In this chapter, the boarding school epidemic will be investigated again, with the stochastic Galerkin method being used to make predictions throughout the epidemic. Only data that would have been available at that time will be used and future data points will be ignored.

Roberts [12] used the stochastic Galerkin method to predict the likely course of an influenza outbreak in New Zealand using only data that would have been available during the epidemic. However, only one prediction was made and the predictions were not updated using later data points. Uncertainty was only considered in the value of $R_0$, and it was assumed to be uniformly distributed.

When applying the stochastic Galerkin method to make predictions throughout the boarding school epidemic, uncertainty will be considered in both $\beta$ and $\gamma$. Also, similar to Chapter 5, the probability distributions of the uncertain parameters will be calculated directly from the data.
Throughout this chapter, it is important to remember that the final outcome of the epidemic is known as it was discussed in detail in Chapters 5 and 6. Therefore, while trying to make predictions in this chapter that could have occurred during the epidemic, it is important to not use any conclusions, such as possible parameter values or ranges, from work that was included earlier in this thesis.

7.1 Using a ‘best fit’ approach

The first step in predicting the likely course the boarding school epidemic would take is choosing a suitable epidemic model. An SIR model is the simplest of the epidemic models that accounts for recovery, but an SEIR model to account for a latent period would also be appropriate. Also, at the beginning of the epidemic, it would have been unknown if reinfection was possible. Therefore SIRS or SEIRS models could also have been chosen. For simplicity, an SIR model will be used in this chapter as it only has two uncertain parameters. However, it is important to note that it is known that no reinfection occurred in the boarding school and that, in Chapter 5, it was shown that an SEIR model provided no benefit over an SIR model. Therefore, it is possible that the decision of ‘simplicity’ is swayed by other factors.

At the start of the epidemic, it was known that there was initially one student infected. Additionally, it is assumed that the rest of the students are susceptible to the disease. Therefore, the initial conditions of the SIR model are known with certainty. It is only the parameters $\beta$ and $\gamma$ that need to be determined and contain uncertainty.

One of the simplest ways of predicting the likely course the boarding school epidemic would take is to use a least squares fitting technique to find the ‘best’ values for the parameters in the SIR model. For chosen values of $\beta$ and $\gamma$, the error in the prediction made $D$ days after the start of the epidemic, $E_{\beta,\gamma}^D$, is given by

$$E_{\beta,\gamma}^D = \sqrt{\sum_{k=1}^{D} [I_{\beta,\gamma}(k) - I_R(k)]^2}$$  \hspace{1cm} (7.1)

where $I_{\beta,\gamma}(k)$ is the fraction of individuals infected on day $k$ according to the SIR model and $I_R(k)$ is the recorded fraction of individuals infected on day $k$. Note that this error formula is similar to the error formula used in Chapters 5 (Equation (5.1)). However, in Equation (7.1), the summation only goes up to the day, $D$, that the prediction is made, as all future data points would be unknown on that day. Therefore, only a subset of the data in Table 5.1 is used, whereas Equation (5.1), which is used in Chapters 5 and 6, uses all fifteen data points.

Unfortunately, while the ‘best fit’ method gives the lowest error, there is no
Chapter 7. Using the stochastic Galerkin method as a predictive tool

Figure 7.1 shows a ‘best fit’ prediction made on day four of the epidemic. The ‘best fit’ prediction has parameter values of $\beta \approx 1.0924$ and $\gamma = 0$ with an error of approximately 0.0025. The parameter values were found using the MATLAB function `fminsearch`. The ‘best fit’ prediction passes through the known data points extremely well. However, it does not predict the peak or tail of the epidemic very well at all. With a $\gamma$ value of zero, the ‘best fit’ is essentially an $SI$ model. Assuming $I_{\text{initial}} > 0$, the long term behaviour of an $SI$ model always predicts the entire population becoming infected. This is because an $SI$ model, or a $\gamma$ value of 0 in an $SIR$ model, assumes that an infected person cannot recover from the disease. This is clearly different to what actually happens and is an unrealistic prediction.

It is important to note that when using `fminsearch`, the parameter values were forced to be non-negative. If this condition was not imposed, the parameter values with the lowest error were $\beta \approx 0.4913$ and $\gamma \approx -0.5828$, which gave an error of approximately 0.0020. However, as $1/\gamma$ is the average recovery rate, it does not make sense for $\gamma$ to be negative. In fact, a negative value of $\gamma$ causes the fraction of infected individuals, $I$, to grow well above 1, which is not physically possible.
7.2 Considering a range of plausible parameter values

Rather than just considering the ‘best’ values that result in the lowest error to make a prediction, a range of plausible values for the parameters is instead considered that keep the error low (similar to Chapter 5). The best fit prediction gives the lowest possible error as well as a good indication of what would be considered to be a low error. Therefore two error thresholds were considered, double and triple the ‘best fit’ error.

7.2.1 Day four predictions

As the ‘best fit’ prediction on day four had an error of approximately 0.0025, the two error thresholds that will be considered when determining probability distributions for $\beta$ and $\gamma$ for the stochastic Galerkin method are 0.005 and 0.0075.

Figure 7.2 shows a heat map of the error on day four of the epidemic, as well as contours of plausible parameter values based upon the two error thresholds. As the two contours are very narrow and close together, the heat map was difficult to produce and not as clear as desired. From the heat map, it can be seen that for the larger error threshold, plausible $\beta$ values range from approximately 1 to 2.4 while plausible $\gamma$ values range from 0 to approximately 1.3. Therefore the plausible recovery times can vary between not being able to recover from the disease ($\gamma = 0$), to being able to recover in less than a day ($1/1.3 \approx 0.77$ days). Even when care was taken to ensure that the error is kept close to its lowest value, the parameters have quite a large range which could lead to extremely different predictions. This is why the ‘best fit’ prediction can be very different from what actually happens, which can be clearly seen in Figure 7.1.

From the error heat map, probability distributions for $\beta$ and $\gamma$ can be determined using a similar method to Chapter 5. Figure 7.3 shows the probability distributions for $\beta$ and $\gamma$ for predictions made on day four of the epidemic. As was seen from the heat map, the probability distributions are quite wide for both of parameters, even when using the smaller of the two error thresholds. Also, because the contours on the heat map are so narrow, the probability distributions are not smooth. For example, the $\beta$ distribution for the lower error bound (red ×) initially changes between five and six corresponding $\gamma$ values within the contour. By only changing by one corresponding $\gamma$ value, this means the $\beta$ values all have approximately the same probability. However, as the contours are so narrow, the difference of one corresponding $\gamma$ value looks significant on the plots of the probability distributions and makes it look as if the probability distributions are highly oscillatory, which is not the case. Because of this, the probability distributions, especially for the lower error threshold, have been modelled rather poorly by the fifth order polynomials.
While the polynomials poorly approximated the probability distribution, it still gave an analytic form for the probability density functions of the uncertain parameters. The associated orthogonal polynomials were then calculated using the Gram-Schmidt orthonormalisation method. The stochastic Galerkin method was then applied to calculate the mean predictions based upon the two error thresholds, as well as their variances.

Figure 7.4 shows the predictions made on day four of the epidemic using both the ‘best fit’ and stochastic Galerkin methods. Unfortunately, using the larger error threshold, the stochastic Galerkin solution blew-out and so its prediction is not useful. However, using the lower error threshold, a prediction could be calculated as well as its variance. The stochastic Galerkin prediction slightly overestimates the peak of the epidemic and models the tail of the epidemic rather poorly. However, it provides a much better prediction than the ‘best fit’ method. The mean prediction also provides a good indication of what the peak of the epidemic could possibly be, as well as when it will occur. Also, with the exception of the last three data points, all data points fall within one standard deviation of the mean prediction. Although, it must be noted that the standard deviation is quite large and, therefore, it is not surprising that it would encompass the data points.

It has been noted by researchers that when making predictions during an epidemic, the timing of the peak of the epidemic is easier to forecast than the height of the peak and often results in the peak being overestimated or underestimated [78, 80]. The peak of the epidemic has been predicted a day late, but this is reasonable considering so few data points are available on day four of the epidemic.

### 7.2.2 Day five predictions

Using a similar methodology to the day four predictions, updated predictions can be calculated on day five of the epidemic. This gives an additional data point to use in the calculations. The ‘best fit’ on day five has an error of approximately 0.02. While this seems like a very small error, it is approximately eight times larger than the ‘best fit’ error on day four of the epidemic. As the ‘best fit’ error is approximately 0.02, the two error thresholds that will be considered for the stochastic Galerkin predictions are 0.04 and 0.06.

Figure 7.5 shows a heat map of the error as well as the contours of plausible parameter values based upon the two error thresholds. From the heat map, it can be seen that the range of plausible \( \beta \) and \( \gamma \) values has narrowed compared to plausible ranges calculated on day four of the epidemic. Unfortunately, the range of plausible \( \gamma \) values still results in very different predictions. As \( \gamma = 0 \) is still a plausible value, a possible prediction is everyone becoming infected with no one able to recover. However, the maximum \( \gamma \) value has reduced from approximately 1.3 to approximately 0.7. While this is still a very large range of \( \gamma \) values, it is hoped that
Figure 7.2: Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 4 of the epidemic. Inner contour has an error threshold of 0.0050 and the outer contour has an error of 0.0075. The two contours are very narrow and very close together. Narrowing the range of parameter values will lead to more accurate predictions.

Figure 7.6 shows the probability distributions for $\beta$ and $\gamma$ determined from the error heat map. While the contours on the error heat map are still quite narrow, the probability distributions on day five are much smoother than those determined on day four. By widening the contours, for each plausible $\beta$ value, there are more corresponding $\gamma$ values. Therefore, the difference of one corresponding $\gamma$ value has far less of an effect on the probability distribution that it did on day four. Because of this, the fifth order polynomials approximate the distributions much better than on day four.

Figure 7.3: Plausible $\beta$ and $\gamma$ probability distributions on day 4 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively.
Chapter 7. Using the stochastic Galerkin method as a predictive tool

Figure 7.4: Predictions made on day 4 of the epidemic. Squares are known data points, while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean.

Figure 7.7 shows the different predictions made on day five of the epidemic. The ‘best fit’ prediction is still predicting that the entire population will be infected (as $\gamma = 0$). Both of the stochastic Galerkin predictions were able to be calculated and did not blow-out. The peak of the epidemic was predicated two days late using the lower error threshold but only one day late using the larger error threshold. Unfortunately, the means of both stochastic Galerkin predictions overestimate the peak of the epidemic as well as its tail. Using the lower error threshold, the mean peak of the epidemic is predicted to have over 60% of the students infected, when the actual peak only had 37% of the students infected. The lower error threshold also predicted that the peak of the epidemic would occur approximately two days later than it did. The larger error threshold predicts the peak of the epidemic much better. It predicted that at the epidemic’s peak, approximately 45% of the students would be infected. The peak also occurs only slightly later than the actual peak.

Using the lower error threshold, unfortunately only one future data point falls within one standard deviation of the mean prediction. However, using the larger error threshold, most future data points fall within the confidence interval. This is a surprising result as it was expected that the lower error threshold would more accurately predict the epidemic. However, using a larger error threshold allowed for parameter values that more accurately predicted the epidemic to be included and therefore provided better predictions.
Chapter 7. Using the stochastic Galerkin method as a predictive tool

7.2.3 Day six predictions

New predictions were calculated on day six of the epidemic using the new data point. Figure 7.8 shows a heat map of the error on day six of the epidemic as well as contours of the plausible parameter values. It is interesting to note that using the lower error threshold, $\gamma = 0$ is no longer a plausible value and now has a minimum value slightly above 0.1. The contours are also much wider when compared to the day four and five contours.

Figure 7.9 shows the probability distributions of the parameters determined from the error heat map. The curves are now much smoother due to the widening of the
Figure 7.7: Predictions made on day 5 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean.

contours on the heat map. This makes it so that the polynomials nicely approximate the probability distributions.

Figure 7.10 shows the different predictions made on day six of the epidemic. Most notably, the ‘best fit’ prediction is much better on day six compared to earlier predictions. It no longer predicts the entire population becoming infected and models the peak of the epidemic rather well, but still slightly overestimates it. The stochastic Galerkin predictions have approximately the same peak as the ‘best fit’ prediction, but all three approximate the tail of the epidemic differently, with the ‘best fit’ approximating the tail the best. Using the lower error threshold, all but the last three data points fall within one standard deviation of the mean and using the higher error threshold, all data points fall within the confidence interval. As all three predictions are quite similar, this gives confidence that the peak of the epidemic has been modelled correctly. On previous days, the particular prediction that is most likely to be accurate would need to be determined.

### 7.2.4 Day seven predictions

On day seven, the fraction of infected individuals decreased for the first time in the epidemic. This is a very important data point as it shows the peak of the epidemic, as well as giving an indication as to the initial slope of the decline of the epidemic.
Figure 7.8: Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 6 of the epidemic.

Figure 7.11 shows a heat map of the error as well as contours of the plausible parameter values. Even using the larger error threshold, the minimum $\gamma$ value is approximately 0.3 whereas on day six, the minimum plausible $\gamma$ value was still zero. The contours have once again become wider while the range of plausible parameter values has narrowed. Figure 7.12 shows the probability distributions of the parameters. Like the day six distributions, the day seven distributions are quite smooth and can be accurately approximated with polynomials.

Figure 7.13 shows the different predictions made on day seven of the epidemic. All three predictions are very similar with the ‘best fit’ having the highest peak. Both stochastic Galerkin predictions slightly underestimate the peak of the epidemic, but

Figure 7.9: Plausible $\beta$ and $\gamma$ probability distributions on day 6 of the epidemic determined from the error heat map. Red is double and blue is triple the ‘best fit’ error, respectively.
accurately predict when it will occur. All three predictions underestimate the frac-
tion infected on days eight, nine and ten, while overestimating the fraction infected
on days twelve through fourteen. The confidence intervals have significantly reduced
in size with most of the data points falling within the confidence intervals.

7.2.5 Day nine predictions

On day nine of the epidemic, the number of infected students once again decreased.
This confirms that the epidemic has passed its peak and is in decline. Figure 7.14
shows a heat map of the error for the parameters on day nine of the epidemic.
Comparing this heat map to the heat map generated on day seven (Figure 7.11),
the range of plausible values for both $\beta$ and $\gamma$ on day nine has decreased slightly
compared to day seven. For example, on day seven the maximum $\gamma$ value was
approximately 0.65 whereas on day nine this has decreased to slightly below 0.6.
Comparing the probability distributions (Figures 7.12 and 7.15), it can be seen that
the probability distributions have slightly narrowed.

Figure 7.16 shows the different predictions made on day nine of the epidemic. The
predictions are very similar to those made on day seven. The tail of the predictions
is slightly higher, better approximating the fraction infected on days eight through
ten. The predictions still overestimate the fraction infected on days twelve through
It is interesting to note that the confidence intervals for the tail of the epidemic are much narrower than at the peak of the epidemic. While the peak of the epidemic has already passed, the confidence intervals are wider than for future days where no data is known.

### 7.2.6 Predictions made after day nine

Predictions made after day nine were very similar to the predictions made on day nine. This is because the peak of the epidemic had already passed and the slope of the tail had already been established. Figure 7.17 shows a heat map of the error on
Figure 7.13: Predictions made on day 7 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean.

Figure 7.14: Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 9 of the epidemic.
day eleven of the epidemic. Compared to the heat map on day nine (Figure 7.14), the contours are very similar. This leads to very similar probability distributions (Figures 7.15 and 7.18) and very similar predictions (Figures 7.16 and 7.19). By day eleven, the epidemic is mostly over and further predictions are unlikely to helpful.

7.3 Possible improvements to early predictions

Early in the epidemic, it was seen that even when the error was kept low (close to the ‘best fit’ error), the range of plausible parameter values could be quite large. This led to very different predictions, as well as to the stochastic Galerkin predictions having a large variance. This was especially apparent with the large range in γ values. When making predictions on day four, plausible values of γ ranged from 0 to 1.3. While this may seem like a small range, this range of γ values corresponds to being able to recover from the disease in less than a day to never recovering from the disease and remaining infectious. These are extremely different real world scenarios.

Rather than simply finding plausible parameter values from the known data points alone, information obtained about the disease could be used to restrict the ranges of the parameters. For example, as it was an influenza outbreak, minimum and maximum values of γ could be enforced based upon known minimum and maximum recovery times. Simple assumptions such as a minimum recovery time of a day and a maximum recovery time of a week (1/7 ≤ γ ≤ 1) would have significantly reduced the range of plausible γ values. As students recovered from the disease, data about length of infection could have been collected. Especially early in the epidemic, this would have been extremely helpful in determining plausible ranges of γ. By restricting the range of plausible γ values, this would have also restricted the range of β values. With narrower distributions, more accurate predictions could hopefully be obtained early in the epidemic.

If the disease was to spread outside the boarding school, before the end of the epidemic within the boarding school, parameter probability distributions from the
Figure 7.16: Predictions made on day 9 of the epidemic. Squares are known data points while circles are future data points. Black is the ‘best fit’ prediction. Red and blue are stochastic Galerkin predictions with double and triple the error of the ‘best fit’, respectively. Solid lines are mean solutions and dashed lines are one standard deviation from the mean.

Figure 7.17: Error heat map showing ranges of plausible $\beta$ and $\gamma$ values on day 11 of the epidemic.
boarding school could be used to help predict the likely course of the epidemic outside the boarding school. The boarding school would likely be several days further through the epidemic as it originated there. Therefore the probability distributions from the boarding school are likely to be more accurate due to having more data points. Each day, when new probability distributions are calculated to update predictions for the boarding school epidemic, the probability distributions could also be used to calculated updated predictions for outside the boarding school. This could potentially give much better predictions than using only data from outside the boarding school.
Chapter 8

Conclusion and future work

Infectious diseases are a serious problem throughout the world. The ability to mathematically model the spread of the disease is of great importance as model predictions can help prepare for the coming epidemic as well as predict the effectiveness of possible interventions. While compartmental epidemic models have been widely studied and used to make predictions, the parameters within these models, as well as their initial conditions, are often not known with certainty.

Uncertainty in the parameters and initial conditions of a compartmental epidemic model can be represented using random variables with known probability distributions. However, due to the presence of the random variables in the epidemic model, it can no longer be solved using a single call to an ODE solver. Sampling techniques, such as Monte Carlo sampling, can be used to determine the mean solution and its variance. One advantage of sampling techniques is that they can usually utilise existing code that has been written for the deterministic model. However, the model needs to be solved numerous times, which can be computationally expensive.

This thesis investigated applications of the stochastic Galerkin method to compartmental epidemic models with uncertainty in the parameters and initial conditions. The stochastic Galerkin method employs a spectral approach, using orthogonal polynomials as basis functions. Applying the stochastic Galerkin method eliminates the random variables and results in a larger system of deterministic ODEs. As the final equations are deterministic, they need only be solved once. The mean prediction and its variance can be obtained directly from the stochastic Galerkin expansion. This provides a significant speed increase over sampling techniques.

In this thesis, the stochastic Galerkin method was applied to an SIR model with different combinations of uncertain parameters and initial conditions. Even when using low order expansions, the stochastic Galerkin method was shown to produce accurate results when compared to results obtained through Monte Carlo sampling.

One disadvantage of the stochastic Galerkin method is that, unlike sampling techniques, it is not able to utilise existing deterministic code for the particular model. Instead, new code is usually required to symbolically calculate the final
system of deterministic ODEs. While this is not a problem for simple compartmental models such as the SIS and SIR used throughout this thesis, this is undesirable for large, complex models [19].

This thesis showed that it is possible for blow-outs to occur when applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters. The blow-outs resulted in the solutions for $S_i(t)$ and $I_i(t)$ to grow to large, unreasonable values. When this occurred, neither the mean prediction, nor its variance, could be calculated. This is arguably the biggest drawback to using the stochastic Galerkin method. Some researchers had previously noted problems when the uncertain parameters had normal distributions [12, 13]. However, they quickly dismissed this as being due to the parameters taking unrealistic (negative) values, since a normally distributed random variable can take any real value. However, this thesis has shown that blow-outs are not isolated to when the uncertain parameters are normally distributed. Blow-outs can occur even when the parameters are gamma distributed, and hence are always positive.

As blow-outs had previously not been studied, despite their obvious drawbacks, this thesis explored blow-outs in the the stochastic Galerkin method. Using a low-order stochastic Galerkin expansion to an SIS model, it was found that, if the final system of deterministic differential equations lacked a stable attractor, the stochastic Galerkin method would blow-out. It was also found that for low order expansions, the range of blow-outs could be calculated analytically using the discriminant. It was shown that the range of blow-outs calculated analytically agreed very well with what was found numerically. While ranges of blow-outs could be calculated analytically for low order expansions, this was not possible for higher order expansions and instead the range of blow-outs could be found only numerically.

While the lack of a stable attractor was the most common reason for a blow-out, it was also found that the initial conditions could also cause a blow-out, despite the presence of a stable attractor. This would occur when the location of the stable or unstable manifold did not allow the solution from particular initial conditions to reach the stable attractor. While this thesis significantly furthered the understanding of the blow-outs that can occur when using the stochastic Galerkin method, more research is needed in this area.

As an application of the methodology studied in this thesis, an epidemic that went through a small boarding school was then investigated. Rather than simply calculating the parameters that resulted in the most accurate fit to the data, a range of parameter values was instead considered that kept the resulting fit reasonably accurate. Initially this was done using approximate Bayesian computation, but since this required the model to be solved a very large number of times to obtain reasonably accurate probability distributions for the uncertain parameters, it was decided that this method was undesirable. An alternative method was proposed
based on the finding that the range of plausible values formed a simple closed shape on a 2D plot. Using this closed shape, probability distributions for the uncertain parameters could be determined. This significantly reduced the number of times the model needed to be solved compared to approximate Bayesian computation. A further reduction in computation time was achieved from the observation that only the border of the closed shape needed to be found in order to determine probability distributions for the uncertain parameters.

The stochastic Galerkin method was extended to work with uncertain parameters of arbitrary probability distributions. As the probability distributions for the uncertain parameters were non-standard (not uniform, gamma, beta or normal distributions), the probability distributions were approximated with polynomials. The associated orthogonal polynomials were then derived using a Gram-Schmidt orthonormalisation method and the stochastic Galerkin method was applied to find the mean prediction and its variance. Different error thresholds for the uncertain parameters led to different predictions with higher error thresholds having larger variances.

The boarding house epidemic was then revisited, under the assumption that the uncertain parameters were no longer independent. A dependent distribution for the parameters was found by plotting 1/error. This caused parameter pairs with a high error to have low probabilities, while parameter pairs with a low error had higher probabilities. The surface was then approximated with exponential functions and normalised by dividing by the volume under the surface. Orthogonal polynomials were derived such that their weight function matched the probability density function of the uncertain parameters. Unfortunately, the predictions obtained were less accurate than those obtained when the uncertain parameters were assumed to have dependent distributions. While the predictions were not as accurate as anticipated, it is hoped that this method will be helpful in the future. Different approaches to determining the dependent probability distribution of the parameters could possibly lead to more accurate predictions.

Finally, the methods used earlier in the thesis were adapted to be used as a predictive tool during an epidemic. On each day of the boarding school epidemic, probability distributions for the uncertain parameters were determined. The associated orthogonal polynomials were derived and the stochastic Galerkin method applied to obtain the mean predictions and their variance. Predictions made early in the epidemic were the least accurate, as the range of plausible values was quite large. However, by using the stochastic Galerkin method, estimates of the peak of the epidemic could be determined whereas the ‘best fit’ approach simply predicted the entire school becoming infected. Predictions made after the peak of the epidemic were quite accurate, due to the narrowing of the range of plausible parameter values.

While there had already been some research into applications of the stochas-
tic Galerkin method to epidemic models with uncertainty in their parameters, this thesis provides several significant extensions to the work previously published, including using arbitrary distributions for the uncertain parameters and investigating the cause of blow-outs in the stochastic Galerkin method.

8.1 Future work

While this thesis significantly expands the understanding and applications of the stochastic Galerkin method to epidemic models with uncertainty in their parameters, this section details some possible avenues of future research that time did not allow to be completed and included in this thesis.

Arguably the greatest need for future research relates to the blow-outs that can occur when applying the stochastic Galerkin method to epidemic models with uncertainty in their parameters and initial conditions. The work conducted while writing this thesis concluded that the cause of blow-outs is usually due to the absence of a stable attractor, but that this is not always the case. Ideally, a ‘work-around’ or modification to the stochastic Galerkin method could be found that avoids the possibility of blow-outs altogether. If this is not possible, it would be advantageous to have a simple ‘test’ that determines if the stochastic Galerkin method is likely to blow-out before starting to apply the method. Deriving the system of deterministic ODEs is essentially wasted time if a different method ultimately must be used.

The methodology associated with the boarding house predictions could possibly also be extended. Throughout this thesis, as an SIR model was used, there were only two uncertain parameters. If a different epidemic model were used, with more uncertain parameters and initial conditions, the plausible range of parameter values could no longer be shown on a 2D plot. This thesis introduced the idea of finding the border of the plausible values, which significantly decreased the number of parameter pairs tested. However with more than two parameters, this method in its current form would not be suitable, and an extension would need to be developed.

Finally, applications of the stochastic collocation method [63, 19] to epidemic models with uncertainty in their parameters would be a logical area of research to partner with the research already presented in this thesis. The stochastic collocation method is an extension to deterministic collocation techniques such as Gaussian quadrature [86], and as with the stochastic Galerkin method, is based on the generalised polynomial expansion. However, unlike the stochastic Galerkin method, the deterministic functions are not determined. Instead, the model is evaluated at appropriately chosen values of the uncertain parameters, called nodes, which are the roots of the associated orthogonal polynomials [87]. The mean and variance of the model can then be determined by multiplying the model evaluations at the nodes with predetermined weights and summing them together [88]. Additional accuracy
can be obtained by increasing the number of nodes.

In addition to being significantly faster than Monte Carlo sampling, one significant advantage of the stochastic collocation method is that it can utilise existing deterministic code for the model, since the stochastic collocation solution simply relies on the model’s evaluation at the nodes [63]. Additionally, there is no indication, as yet, that blow-outs occur when implementing the stochastic collocation method. This makes it a logical alternative for regions in which the stochastic Galerkin method blows-out.
Appendix A

Below is the MATLAB routine to calculate the deterministic system of ODEs that result from the stochastic Galerkin method, for an SIR model with uncertainty in \( \beta \) and \( \gamma \). See Section 3.6 and Equation 3.29 for additional details on the application of the stochastic Galerkin method to the SIR model.

The routine is currently configured for uniformly distributed \( \beta \) and \( \gamma \), using the Legendre polynomials. However, the code can quickly be adapted to other probability distributions for \( \beta \) and \( \gamma \). The functions `legendre_poly3` and `legendre_poly4` simply need to be replaced with functions that calculate the integrals of the appropriate orthogonal polynomials.

For convenience, the integrals of the Legendre polynomials are calculated symbolically. However, as many of the integrals are used several times while symbolically generating the deterministic ODEs, it would be better to precalculate the integrals and store the results in an array. For low order expansions, this is not necessary. However, this provides a significant speed increase for higher order expansions.

```matlab
function SIR_uniform_equations(P)
% Calculates the system of deterministic ODEs for the SIR epidemic model
% with uncertainty in \beta and \gamma. Both \beta and \gamma are uniformly
% distributed. See Equation 3.29 on page 56.
% P is the order of the stochastic Galerkin expansion

fprintf('
')
for u = 0:1:P
    for v = 0:1:(P-u)
        K_uv = 1/((2*u+1)*(2*v+1));
        fprintf('dS_%i%i/dt = - beta_0/%d * (0',u,v,K_uv)
        for i = 0:1:P
            for j = 0:1:(P-i)
                for m = 0:1:P
                    for n = 0:1:(P-m)
                        integral = legendre_poly3(i,m,u);
                        if integral ~= 0
                            integral = integral * legendre_poly3(j,n,v);
                            if integral ~= 0
                                if integral == 1
                                    fprintf('+ S_%i%i*I_%i%i',i,j,m,n)
                                else
                                    fprintf('+ %d*S_%i%i*I_%i%i',integral,i,j,m,n)
                                end
                            end
                        end
                    end
                end
            end
        end
    end
end
```
fprintf(') - a/%d*(0',K_uv)
for i = 0:1:P
    for j = 0:1:(P-i)
        for m = 0:1:P
            for n = 0:1:(P-m)
                integral = legendre_poly4(1,i,m,u);
                if integral ~= 0
                    integral = integral * legendre_poly3(j,n,v);
                    if integral ~= 0
                        if integral == 1
                            fprintf('+ S_%i%i*I_%i%i',i,j,m,n)
                        else
                            fprintf('+ %d*S_%i%i*I_%i%i',integral,i,j,m,n)
                        end
                    end
                end
            end
        end
    end
end
fprintf('); 
')

dI_%i%i/dt = beta_0/%d * (0',u,v,K_uv)
for i = 0:1:P
    for j = 0:1:(P-i)
        for m = 0:1:P
            for n = 0:1:(P-m)
                integral = legendre_poly3(i,m,u);
                if integral ~= 0
                    integral = integral * legendre_poly3(j,n,v);
                    if integral ~= 0
                        if integral == 1
                            fprintf('+ S_%i%i*I_%i%i',i,j,m,n)
                        else
                            fprintf('+ %d*S_%i%i*I_%i%i',integral,i,j,m,n)
                        end
                    end
                end
            end
        end
    end
end
fprintf(') + a/%d*(0',K_uv)
for i = 0:1:P
    for j = 0:1:(P-i)
        for m = 0:1:P
            for n = 0:1:(P-m)
                integral = legendre_poly4(1,i,m,u);
                if integral ~= 0
                    integral = integral * legendre_poly3(j,n,v);
                    if integral ~= 0
                        if integral == 1
                            fprintf('+ S_%i%i*I_%i%i',i,j,m,n)
                        else
                            fprintf('+ %d*S_%i%i*I_%i%i',integral,i,j,m,n)
                        end
                    end
                end
            end
        end
    end
end
fprintf('); 
')
fprintf(') - gamma_0*I_%i%i - b/%d*(0',u,v,K_uv)
for i = 0:1:P
  for j = 0:1:(P-i)
    if i == u
      integral = 1/(2*u+1) * legendre_poly3(i,j,v);
      if integral ~= 0
        if integral == 1
          fprintf('+ I_%i%i',i,j)
        else
          fprintf('+ %d*I_%i%i',integral,i,j)
        end
      end
    end
  end
end
fprintf('); 
')
end

function integral = legendre_poly3(i,j,k)
% Calculates the integral of the product of three Legendre polynomials.
% Note that P(1) is the zeroth order Legendre polynomial and P(2) is the
% first order Legendre polynomial etc.
syms x;
P{1} = 1;
P{2} = x;
P{3} = (3*x^2 - 1)/2;
P{4} = (5*x^3)/2 - (3*x)/2;
P{5} = (35*x^4)/8 - (15*x^2)/4 + 3/8;
P{6} = (63*x^5)/8 - (35*x^3)/4 + (15*x)/8;
P{7} = (231*x^6)/16 - (315*x^4)/16 + (105*x^2)/16 - 5/16;
P{8} = (429*x^7)/16 - (693*x^5)/16 + (315*x^3)/16 - (35*x)/16;
P{9} = (6435*x^8)/128 - (3003*x^6)/32 + (3465*x^4)/64 - (315*x^2)/32 + 35/128;
P{10} = (12155*x^9)/128 - (6435*x^7)/32 + (9009*x^5)/64 - (1155*x^3)/32 + (315*x)/128;
P{11} = (46189*x^10)/256 - (109395*x^8)/256 + (45045*x^6)/128 - (15015*x^4)/128 + (3465*x^2)/256 - 63/256;
integral = double(int(P{i+1} * P{j+1} * P{k+1} * 1/2,x,-1,1));
end

function integral = legendre_poly4(i,j,m,n)
% Calculates the integral of the product of four Legendre polynomials.
% Note that P(1) is the zeroth order Legendre polynomial and P(2) is the
% first order Legendre polynomial etc.
syms x;
P{1} = 1;
P{2} = x;
P{3} = (3*x^2 - 1)/2;
P{4} = (5*x^3)/2 - (3*x)/2;
P{5} = (35*x^4)/8 - (15*x^2)/4 + 3/8;
P{6} = (63*x^5)/8 - (35*x^3)/4 + (15*x)/8;
P{7} = (231*x^6)/16 - (315*x^4)/16 + (105*x^2)/16 - 5/16;
P{8} = (429*x^7)/16 - (693*x^5)/16 + (315*x^3)/16 - (35*x)/16;
P{9} = (6435*x^8)/128 - (3003*x^6)/32 + (3465*x^4)/64 - (315*x^2)/32 + 35/128;
P{10} = (12155*x^9)/128 - (6435*x^7)/32 + (9009*x^5)/64 - (1155*x^3)/32 + (315*x)/128;
P{11} = (46189*x^10)/256 - (109395*x^8)/256 + (45045*x^6)/128 - (15015*x^4)/128 + (3465*x^2)/256 - 63/256;
integral = double(int(P{i+1} * P{j+1} * P{m+1} * P{n+1} * 1/2,x,-1,1));
end
Appendix B

Below is the MATLAB routine to find the border of the closed shape of plausible parameter values for the boarding school epidemic. By only finding the border, many parameter pairs do not need to be tested, which results in a significant speed increase. The probability distributions of the uncertain parameters, $\beta$ and $\gamma$, are plotted and approximated by fifth order polynomials using the inbuilt function *polyfit*.

```matlab
function [beta_actual_min,beta_actual_max,gamma_actual_min,gamma_actual_max, ... 
    beta_polynomial,gamma_polynomial] = find_border

% Finds the border of the closed shape of plausible values for the boarding 
% school epidemic data and uses it to find the probability distributions 
% for beta and gamma.  
% % beta_actual_min and beta_actual_max are the minimum and maximum plausible 
% % values for beta.  
% % gamma_actual_min and gamma_actual_max are the minimum and maximum 
% % plausible values for gamma.  
% % beta_polynomial and gamma_polynomials are the polynomials that 
% % approximate the beta and gamma probability distributions, respectively 
% % 0 is false, 1 is true

close all;  
% error threshold
max_error = 0.25;

% figure for closed shape
figure(1)
hold on;
xlabel('$\beta$', 'Interpreter', 'LaTeX')
ylabel('$\gamma$', 'Interpreter', 'LaTeX')

% beta probability scatter
figure(2)
hold on;
xlabel('$\beta$', 'Interpreter', 'LaTeX')
ylabel('$P(\beta)$', 'Interpreter', 'LaTeX')

% gamma probability scatter
figure(3)
hold on;
xlabel('$\gamma$', 'Interpreter', 'LaTeX')
ylabel('$P(\gamma)$', 'Interpreter', 'LaTeX')

% boarding house data
recorded_data = [1,3,7,25,72,222,282,256,233,189,123,70,25,11,4];
N = 763;
normalised_data = recorded_data/N;
```

172
% minimum, maximum and starting values for beta and gamma
beta_min = 1.2;
beta_max = 2.2;
beta_step = 0.001;
beta_start = 1.7;
gamma_min = 0.2;
gamma_max = 0.9;
gamma_step = 0.001;
gamma_start = 0.5;

% initialising actual minimum and maximum values of beta found that are
% below the error threshold
beta_actual_max = 0;
beta_actual_min = 0;

% array containing the minimum and maximum plausible gamma values for each
% plausible beta value
% -1 means that it hasn't been tested
beta_boundaries = -1 * ones( round(beta_max - beta_min)/beta_step + 1,2);

% initialising boundary gamma values for starting beta values
inside_upper = gamma_start;
outside_upper = gamma_max;
inside_lower = gamma_start;
outside_lower = gamma_min;

% initially finding upper border
while (outside_upper - inside_upper) > gamma_step * 1.5
    gamma_test_value = inside_upper + round((outside_upper - inside_upper)/gamma_step/2) * gamma_step;
    if isinside(beta_start,gamma_test_value,max_error,normalised_data) == 1
        figure(1)
        inside_upper = gamma_test_value;
        plot(beta_start,gamma_test_value,'b+')
    else
        figure(1)
        outside_upper = gamma_test_value;
        plot(beta_start,gamma_test_value,'rx')
    end
end

% initially finding lower border
while (inside_lower - outside_lower) > gamma_step * 1.5
    gamma_test_value = inside_lower - round((inside_lower - outside_lower)/gamma_step/2) * gamma_step;
    if isinside(beta_start,gamma_test_value,max_error,normalised_data) == 1
        figure(1)
        inside_lower = gamma_test_value;
        plot(beta_start,gamma_test_value,'b+')
    else
        figure(1)
        outside_lower = gamma_test_value;
        plot(beta_start,gamma_test_value,'rx')
    end
end

% storing boundary for starting beta value
beta_position = round((beta_start - beta_min)/beta_step) + 1;
beta_boundaries(beta_position,1) = inside_lower;
beta_boundaries(beta_position,2) = inside_upper;

% increasing beta by step size
beta_value = beta_start + beta_step;
edge_of_shape_found = 0;
while beta_value <= beta_max && edge_of_shape_found == 0
  % collecting information about border from previous beta value
  beta_position = round((beta_value - beta_min)/beta_step) + 1;
  if beta_value > beta_start
    previous_gamma_lower = beta_boundaries(beta_position-1,1);
    previous_gamma_upper = beta_boundaries(beta_position-1,2);
  else
    previous_gamma_lower = beta_boundaries(beta_position+1,1);
    previous_gamma_upper = beta_boundaries(beta_position+1,2);
  end
  previous_gamma_test_value = previous_gamma_upper;
  previous_gamma_status = isinside(beta_value,previous_gamma_test_value,max_error,normalised_data);
  if previous_gamma_status == 1
    figure(1)
    plot(beta_value,previous_gamma_test_value,'b+')
  else
    previous_gamma_test_value = new_gamma_test_value;
    previous_gamma_status = new_gamma_status;
    figure(1)
    plot(beta_value,new_gamma_test_value,'rx')
  end
  upper_edge_found = 0;
  % find upper border for new beta value using information from previous beta value
  while upper_edge_found == 0 && edge_of_shape_found == 0
    if previous_gamma_status == 1
      new_gamma_test_value = previous_gamma_test_value + gamma_step;
      new_gamma_status = isinside(beta_value,new_gamma_test_value,max_error,normalised_data);
      if new_gamma_status == 0
        upper_edge_found = 1;
        beta_boundaries(beta_position,2) = previous_gamma_test_value;
        figure(1)
        plot(beta_value,new_gamma_test_value,'rx')
      else
        previous_gamma_test_value = new_gamma_test_value;
        previous_gamma_status = new_gamma_status;
        figure(1)
        plot(beta_value,new_gamma_test_value,'b+')
      end
    else
      previous_gamma_test_value = new_gamma_test_value - gamma_step;
      new_gamma_status = isinside(beta_value,new_gamma_test_value,max_error,normalised_data);
      if new_gamma_status == 0
        upper_edge_found = 1;
        beta_boundaries(beta_position,2) = new_gamma_test_value;
        figure(1)
        plot(beta_value,new_gamma_test_value,'b+')
      else
        previous_gamma_test_value = new_gamma_test_value;
        previous_gamma_status = new_gamma_status;
        figure(1)
        plot(beta_value,new_gamma_test_value,'rx')
      end
    end
    if previous_gamma_test_value < previous_gamma_lower
      edge_of_shape_found = 1;
      if beta_value > beta_start
        beta_actual_max = beta_value - beta_step;
      else
        beta_actual_min = beta_value + beta_step;
      end
    end
  end
end
% collecting information about border from previous beta value
previous_gamma_test_value = previous_gamma_lower;
previous_gamma_status = isinside(beta_value,previous_gamma_test_value,max_error,normalised_data);
if previous_gamma_status == 1
    figure(1)
    plot(beta_value,previous_gamma_test_value,'b+')
else
    figure(1)
    plot(beta_value,previous_gamma_test_value,'rx')
end

% find lower border for new beta value using information from previous beta value
lower_edge_found = 0;
while lower_edge_found == 0 && edge_of_shape_found == 0
    if previous_gamma_status == 1
        new_gamma_test_value = previous_gamma_test_value - gamma_step;
        new_gamma_status = isinside(beta_value,new_gamma_test_value,max_error,normalised_data);
        if new_gamma_status == 0
            lower_edge_found = 1;
            beta_boundaries(beta_position,1) = previous_gamma_test_value;
            figure(1)
            plot(beta_value,new_gamma_test_value,'rx')
        else
            previous_gamma_test_value = new_gamma_test_value;
            previous_gamma_status = new_gamma_status;
            figure(1)
            plot(beta_value,new_gamma_test_value,'b+')
        end
    else
        new_gamma_test_value = previous_gamma_test_value + gamma_step;
        new_gamma_status = isinside(beta_value,new_gamma_test_value,max_error,normalised_data);
        if new_gamma_status == 0
            previous_gamma_test_value = new_gamma_test_value;
            previous_gamma_status = new_gamma_status;
            figure(1)
            plot(beta_value,new_gamma_test_value,'rx')
        else
            lower_edge_found = 1;
            beta_boundaries(beta_position,1) = new_gamma_test_value;
            figure(1)
            plot(beta_value,new_gamma_test_value,'b+')
        end
    end
end

% updating beta value
if edge_of_shape_found == 0 && beta_value > beta_start
    beta_value = beta_value + beta_step;
else if edge_of_shape_found == 1 && beta_value > beta_start
    beta_value = beta_start - beta_step;
edgede_of_shape_found = 0;
else if edge_of_shape_found == 0 && beta_value < beta_start
    beta_value = beta_value - beta_step;
don

% initialising arrays for storing beta and gamma distributions
beta_dist_x = beta_actual_min:beta_step:beta_actual_max;
beta_dist = zeros( length(beta_dist_x), 1);
gamma_dist = zeros( round((gamma_max - gamma_min)/gamma_step + 1), 1);

beta_boundaries_position = round((beta_actual_min - beta_min)/beta_step + 1);
beta_dist_position = 1;
gamma_actual_min = gamma_max;
gamma_actual_max = gamma_min;

% determining number of points inside boundary for each beta value
% also determines number of points inside boundary for each gamma value
for beta_value = beta_actual_min:beta_step:beta_actual_max
    gamma_upper = beta_boundaries(beta_boundaries_position,2);
    gamma_lower = beta_boundaries(beta_boundaries_position,1);
    if gamma_lower < gamma_actual_min
        gamma_actual_min = gamma_lower;
    end
    if gamma_upper > gamma_actual_max
        gamma_actual_max = gamma_upper;
    end
    gamma_upper_position = round((gamma_upper - gamma_min)/gamma_step + 1);
    gamma_lower_position = round((gamma_lower - gamma_min)/gamma_step + 1);
    for k = gamma_lower_position:1:gamma_upper_position
        gamma_dist(k) = gamma_dist(k) + 1;
    end
    num_points = round((gamma_upper - gamma_lower)/beta_step + 1);
    beta_dist(beta_dist_position) = num_points;
    beta_boundaries_position = beta_boundaries_position + 1;
    beta_dist_position = beta_dist_position + 1;
end

% normalising beta distribution and plotting
beta_area = trapz(beta_dist_x,beta_dist);
beta_dist = beta_dist / beta_area;
figure(2)
plot(beta_actual_min:beta_step:beta_actual_max,beta_dist,'rx')

% normalising gamma distribution and plotting
gamma_actual_min_pos = round((gamma_actual_min - gamma_min)/gamma_step + 1);
gamma_actual_max_pos = round((gamma_actual_max - gamma_min)/gamma_step + 1);
gamma_dist = gamma_dist(gamma_actual_min_pos:1:gamma_actual_max_pos);
gamma_area = trapz(linspace(gamma_actual_min,gamma_actual_max,length(gamma_dist)),gamma_dist);
gamma_dist = gamma_dist / gamma_area;
figure(3)
plot(linspace(gamma_actual_min,gamma_actual_max,length(gamma_dist)),gamma_dist,'rx')

% finding polynomial fit for beta and gamma distributions
polynomial_order = 5;
beta_polynomial = polyfit((beta_actual_min:beta_step:beta_actual_max),beta_dist',polynomial_order);
gamma_polynomial = polyfit((gamma_actual_min:gamma_step:gamma_actual_max),gamma_dist',polynomial_order);
end

function myboolean = isinside(beta_value,gamma_value,max_error,normalised_data)
% isinside determines if the parameter pair, beta_value and gamma_value, have
% an error less than the error threshold
[-y1] = ode45(@SIR_equations,0:1:14,[762/763 1/763],[],beta_value,gamma_value);
error = sqrt(sum((y1(:,2) - normalised_data).^2));
if error < max_error
    myboolean = 1;
else
    myboolean = 0;
end
end
function dydt = SIR_equations(t,y,beta_0, gamma_0)
S = y(1);
I = y(2);
dydt = zeros(2,1);
dydt(1) = -beta_0 * S * I;
dydt(2) = beta_0 * S * I - gamma_0 * I;
end
Bibliography


